BURTON RANDOL

An Introduction to Real Analysis

HARBRACE COLLEGE MATHEMATICS SERIES

An Introduction to Real Analysis

HARBRACE COLLEGE MATHEMATICS SERIES

Salomon Bochner and W. C. Lister, EDITORS

PUBLISHED TITLES

Calculus, Karel de Leeuw
Two-Dimensional Calculus, Robert Osserman
Functions of Several Variables, John W. Woll, Jr.
An Introduction to Real Analysis, Burton Randol
Linear Algebra, Ross A. Beaumont
Algebraic Topology: An Introduction, William S. Massey
Discrete Probability, R. A. Gangolli and Donald Ylvisaker

An Introduction to Real Analysis

BURTON RANDOL
YALE UNIVERSITY

Harcourt, Brace & World, Inc.
New York / Chicago / San Francisco / Atlanta

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 69-10733

PRINTED IN THE UNITED STATES OF AMERICA

Foreword

The Harbrace College Mathematics Series has been undertaken in response to the growing demands for flexibility in college mathematics curricula. This series of concise, single-topic textbooks is designed to serve two primary purposes: First, to provide basic undergraduate text materials in compact, coordinated units. Second, to make available a variety of supplementary textbooks covering single topics.

To carry out these aims, the series editors and the publisher have selected as the foundation of the series a sequence of six textbooks covering functions, calculus, linear algebra, multivariate calculus, theory of functions, and theory of functions of several variables. Complementing this sequence are a number of other planned volumes on such topics as probability, statistics, differential equations, topology, differential geometry, and complex functions.

By permitting more flexibility in the construction of courses and course sequences, this series should encourage diversity and individuality in curricular patterns. Furthermore, if an instructor wishes to devise his own topical sequence for a course, the Harbrace College Mathematics Series provides him with a set of books built around a flexible pattern from which he may choose the elements of his new arrangement. Or, if an instructor wishes to supplement a full-sized textbook, this series provides him with a group of compact treatments of individual topics.

An additional and novel feature of the Harbrace College Mathematics Series is its continuing adaptability. As new topics gain emphasis in the curricula or as promising new treatments appear, books will be added to the series or existing volumes will be revised. In this way, we will meet the changing demands of the instruction of mathematics with both speed and flexibility.

SALOMON BOCHNER

W. G. LISTER

Preface

In this book, which is suitable for a one-semester course in introductory analysis, I have discussed some topics from the theory of functions of one real variable. I have assumed that the reader has had a standard two- or three-semester course in calculus, so that the first two chapters on the real number system and the notion of a function may be, to a certain extent, a remembrance of things past, though perhaps from a more critical viewpoint. More specifically, Chapter One deals with the building up of the real number system from the rational number system by means of Dedekind cuts and also with the theory of sequences and series of real numbers, while Chapter Two deals with the notion of a function and also the properties of various types of functions.

The third and fourth chapters deal, respectively, with power series and the Weierstrass approximation theorem, and some of their applications to analysis, for example, the moment problem and the Karamata Tauberian theorem.

The fifth chapter deals with Fourier series and their applications, such as Hurwitz's approach to the isoperimetric inequality and the introductory theory of equidistributed sequences.

The sixth chapter, on the Lebesgue integral, stands somewhat apart from the rest of the book. It is more abstract than the preceding chapters, and except for some of the problems at the end of the chapter it is presented without applications. This chapter should be regarded as very introductory in flavor. The subject itself is, of course, a very important part of mathematical culture, and the student as he proceeds in mathematics will discover the various reasons for this.

BURTON RANDOL

Contents

CHAPTER ONE

Numbers, Sequences, and Series **1**

1 Introduction 1

2 The inadequacy of the rational number system for certain questions and the development of the real number system 2

3 Infinite sequences and series 8

4 Algebraic operations with series 17

5 Fubini's theorem 21

Exercises 23

CHAPTER TWO

Functions **26**

1 Introductory remarks 26

2 Continuous functions 26

3 Sequences and series of functions 32

4 Other classes of functions 34

Exercises 37

CHAPTER THREE

Power Series **39**

1 Expansions about the origin 39

2 Algebraic operations with power series 44

3 Taylor's theorem 49

4 A C^∞ function that is not analytic 52

5 Expansions about points other than the origin 53

6 A special result 54

7 Concluding remarks 56

Exercises 56

CHAPTER FOUR

The Weierstrass Approximation Theorem **58**

1 The Weierstrass approximation theorem 58

2 The Karamata Tauberian theorem 64

Exercises 69

CHAPTER FIVE

Fourier Series **70**

1 Orthonormal systems 70

2 The notion of completeness 76

3 Some lemmas 77

4 The trigonometric system 80

5 Equidistribution 91

Exercises 94

CHAPTER SIX

The Lebesgue Integral **95**

1 The Lebesgue integral 95

2 Measurable sets 106

Exercises 108

References **109**

Index **111**

CHAPTER ONE

Numbers, Sequences, and Series

1 Introduction

In this chapter, we shall show, among other things, that the rational numbers, or common fractions, whose properties we shall take as known, are by themselves incapable of providing satisfactory answers to certain perfectly natural questions. That is, we shall show that there are properties, that seem certainly to specify some definite number, but to which there cannot correspond any rational number. Confronted with this situation, we can adopt one of two points of view. Either, upon careful scrutiny, the properties in question do not really constitute reasonable requirements for a number, or else the system of rational numbers is not adequate to deal with certain types of questions. Since the questions involved are so interesting, and the psychological evidence so compelling, we shall adopt the second viewpoint. Indeed, if we did not, and admitted only the rationals as numbers, then there would, to give two examples, be no number that could reasonably be regarded as the length of a circle of radius 1, nor would there be any number that could reasonably be regarded as the length of the hypotenuse of an isosceles right triangle having equal sides of length 1. Algebraically, the last example is equivalent, by the Pythagorean theorem, to the impossibility of finding a rational solution to the equation $x^2 = 2$.

Let us examine these difficulties in greater detail. To take matters from the beginning, we recall, very briefly, some facts about the rational numbers. A rational number is an expression of the form p/q, where p and q are both integers, with $q \neq 0$. Apart from this restriction on q, p or q may be positive, negative, or, in the case of p, zero. Recall, also, that the same number can be represented as the quotient of two integers in several ways. For example, $\frac{1}{3} = \frac{2}{6}$, and, in general

$$\frac{p_1}{q_1} = \frac{p_2}{q_2} \qquad \text{if and only if } p_1q_2 = p_2q_1.$$

Because a given rational number can be represented in many ways as the quotient of two integers, it is, strictly speaking, most accurate to describe the rational numbers as the set of "equivalence classes" of expressions of the form of p/q, where each equivalence class consists of those fractions which are equal, in the sense of the above criterion. In practice, however, we can simply refer to "the rational number p/q," without danger of confusion. The reason being that the validity of the usual operations and relationships, such as addition, multiplication, "$<$", etc., does not depend on the particular fractions used to represent the numbers involved. For example, the statements $1 < 2$ and $\frac{9}{9} < \frac{22}{11}$ are different ways of saying the same thing.

We shall, as we have said, take as known the laws governing the various algebraic operations between rationals (addition, multiplication, etc.), and shall, besides, assume that the reader is familiar with the notion of order. That is, between two distinct rationals r_1 and r_2, either $r_1 < r_2$ or $r_2 < r_1$. In terms of particular fractions,

$$\frac{p_1}{q_1} < \frac{p_2}{q_2} \qquad \text{if and only if } p_1q_2 < p_2q_1,$$

provided $q_1q_2 > 0$, while for $q_1q_2 < 0$,

$$\frac{p_1}{q_1} < \frac{p_2}{q_2} \qquad \text{if and only if } p_1q_2 > p_2q_1.$$

2 The inadequacy of the rational number system for certain questions and the development of the real number system

Suppose we take it as reasonably self-evident that there exists some number x, which is the length of the hypotenuse of an isosceles right triangle, whose equal sides have length 1. As we have pointed out this implies that $x^2 = 2$. Now within the system of rationals, we can certainly find numbers that when squared, give results which are arbitrarily close to 2. Thus, for example, $(1.4)^2 = (1 + \frac{4}{10})^2 = 1.96$, $(1.41)^2 = 1.9881$, and still more accurately, $(1.414)^2 = 1.999396$.

However, although progressively more accurate, none of these approximations to our ideal number x has quite the desired property. As it turns out, this is inevitable, since *there is no rational number whose square is 2*. We shall now prove this.

Suppose p/q has the property that $p^2/q^2 = 2$. Without loss of generality, we may assume that the fraction p/q is in lowest terms, that is, p and q have no common factors other than 1 and -1. Now if $p^2/q^2 = 2$, then $p^2 = 2q^2$,

so we conclude that p^2 must be even. Since the square of an odd number is always odd, p itself must be even. This implies that p^2 is divisible by 4. Since $p^2 = 2q^2$ and $p^2/4$ is an integer, we conclude that $q^2/2$ is an integer. This implies that q^2 is even, and hence q itself is even. This, however, is impossible, since the fraction p/q was assumed to be in lowest terms, which implies that p and q cannot both be even. Thus, the desired result is established by contradiction.

This striking result shows in uncompromising terms that if we insist on staying within the framework of the rationals insofar as the concept of a number is concerned, then we must abandon as meaningless the familiar expression $\sqrt{2}$. It is far more profitable to attempt instead to augment the system of rationals in such a way as to give reasonable interpretations to concepts like $\sqrt{2}$.

To see how we might go about this, it is instructive, for purposes of guidance, to imagine the rationals as corresponding, via the familiar graphical interpretation, to certain points on a horizontal straight line infinite in both directions. That is, a point on the line is selected, corresponding to the number 0, and another point, to the right of the first, is chosen, corresponding to the number 1. Once this has been done, each rational number can be made to correspond, in the usual way, to a definite point on the line. For example, the number $\frac{1}{2}$ corresponds to the point midway between the points corresponding to 0 and 1. It should be noted that the points corresponding to the rational numbers are very densely distributed throughout the line. For instance, between the points corresponding to any two distinct rationals r_1 and r_2, there is always a point that corresponds to a third rational. For example, the point which lies midway between the two original points will do, since it corresponds to the rational number $\frac{1}{2}(r_1 + r_2)$, which is the average of r_1 and r_2.

Now as we have seen, there is, insofar as $\sqrt{2}$ is concerned, a definite gap in this system of points, which occurs at the point on the line that lies between the set of points corresponding to the positive rationals whose square is greater than 2, on the one hand, and the remaining rationals, which we shall denote by S, on the other. We are, of course, assuming that such a point exists, but the visual evidence in favor of this is overwhelming, and it is quite evident that it is this point that should correspond, in the graphical representation, to the "number" $\sqrt{2}$. We can, moreover, divorce the concept from any appeal to geometric intuition by simply *defining* $\sqrt{2}$ to be the partition of the rational numbers into S and its complement, or by observing that a partition is completely specified when one is given either of the two sets that define it, we can simply define $\sqrt{2}$ to be the set S.

This simple idea, it turns out, is enormously fruitful, and leads quickly to a satisfactory notion of the system of real numbers. We begin with a definition.

Definition 1.1 By a *Dedekind cut*, or simply a cut, we mean a nonempty subset S of the rationals such that

1. If r_1 is in S, and r_2 is any rational such that $r_2 < r_1$, then r_2 must be in S.
2. S does not contain all the rationals.
3. S contains no greatest element. In other words, there is no rational r in S such that $r' < r$ for every other rational r' in S.

Remark If we assume for a moment that we are already in possession of the real number system, then a cut simply corresponds to the set of all rationals that are less than some given number.

Definition 1.2 By a *real number* we simply mean a Dedekind cut.

This definition, as we have seen, is in harmony with our intuitive notions about the continuity, or absence of gaps, of a straight line. By the *upper section* of a number, or cut, x we mean the set of all rationals not in x.

Since the real numbers are presumably a generalization of the rational numbers, there must be some natural way in which the familiar rationals correspond to certain cuts. The natural correspondence, which we shall adopt, assigns to each rational number r, the set of all rational numbers r' satisfying $r' < r$.

Conversely, we can easily find a criterion to determine whether or not a given cut corresponds in the above way to a rational number. Namely, if x is a cut, and if x' is the upper section of x, then the cut x corresponds to a rational number if and only if x' contains a smallest element. That is, if and only if there exists a rational number r in x' such that $r \leq r'$ for *any* rational r' in x'. The proof of this simple criterion is left to the reader.

We shall, by a harmless abuse of language, call real numbers that correspond in the above way to rationals, *rational numbers*, and we shall call real numbers that are not rational, *irrational numbers*.

Remark It is interesting to note that as the number system has been extended through various stages, the names given to the members of the successive extensions are often of a pejorative character. For example, positive integers, *negative* integers, rational numbers, *irrational* numbers, real numbers, and *imaginary* numbers.

The very important concept of order is easily formulated for real numbers.

Definition 1.3 Given two real numbers x_1 and x_2, we say that $x_1 < x_2$ (or, equivalently, $x_2 > x_1$) if and only if the cut x_1 is a proper subset of the cut x_2. That is, if and only if every element of x_1 is an element of x_2, and there exists some element of x_2 that is not an element of x_1. We say $x_1 = x_2$ if and only if the cuts x_1 and x_2 are the same. The notation $x_1 \leq x_2$ or $x_2 \geq x_1$ means that either $x_1 < x_2$ or $x_1 = x_2$.

It is easy to show that if x_1 and x_2 are real numbers, then precisely one of the following is true:

$$\text{(a) } x_1 = x_2; \qquad \text{(b) } x_1 < x_2; \qquad \text{(c) } x_2 < x_1.$$

To see this, note, to begin with, that it is obviously impossible for more than one of these statements to be true. It remains to be shown that one of them must be true. Suppose $x_1 \neq x_2$. Now either the cut x_1 is a proper subset of the cut x_2 or it is not. In the first event, $x_1 < x_2$. Accordingly, suppose x_1 is not a proper subset of x_2. Then there must exist some rational number r in the cut x_1 that is not in x_2, since we have assumed $x_1 \neq x_2$. Moreover, every rational r' in the cut x_2 must satisfy $r' < r$, since otherwise there would be a rational r'' in x_2 such that $r < r''$, and this would imply, by requirement (1) for cuts, that r is in x_2, which is impossible. Thus, since by requirement (1) for cuts, the cut x_1 must include all rationals up to r, and since we have just shown that all the rationals in x_2 are less than r, we conclude that $x_2 < x_1$.

Definition 1.4 If x is a number, then $-x$ is the cut consisting of all rationals of the form $-r$, where r is in the upper section of x, with the smallest element deleted, if x is rational.

We leave it to the reader to verify that this set satisfies all the requirements for a cut.

Definition 1.5 If x is a number, then $|x|$, the *absolute value* of x, is defined to be

a. x, if $x \geq 0$

b. $-x$, if $x < 0$,

where by 0 we mean the cut that corresponds to the rational number 0.

Definition 1.6 **Addition and Subtraction** If x_1 and x_2 are numbers, then $x_1 + x_2$ is the cut that consists of all rationals of the form $r_1 + r_2$, where r_1 is in the cut x_1, and r_2 is in the cut x_2. By $x_1 - x_2$, we mean $x_1 + (-x_2)$.

The verification that $x_1 + x_2$ is a cut is again left as an exercise.

Definition 1.7 Multiplication Suppose x_1 and x_2 are numbers such that $x_1 > 0$ and $x_2 > 0$. Then x_1x_2 is the cut that consists of the union of the set of all rationals < 0 with the set of all rationals of the form r_1r_2, where r_1 ranges over the rationals that are in the cut x_1 and ≥ 0, and r_2 ranges over the rationals that are in the cut x_2 and ≥ 0. [We again leave to the reader the routine verification that this defines a cut.] If x_1 and x_2 are not both greater than 0, then x_1x_2 is defined to be

a.	0,	if x_1 or x_2 is 0
b.	$\lvert x_1\rvert\,\lvert x_2\rvert$,	if both x_1 and x_2 are less than 0
c.	$-(\lvert x_1\rvert\,\lvert x_2\rvert)$,	if one of the x_i's is greater than 0, and the other is less than 0

Definition 1.8 Division Suppose x is a number such that $x > 0$. Then $1/x$, or x^{-1}, is the cut consisting of the union of the set of all rationals ≤ 0 with the set of all rationals of the form $1/r$, with r in the upper section of x with the smallest element deleted, if x is rational. [As before, we leave to the reader the routine verification that this defines a cut.] If $x < 0$, $1/x$ is defined to be $-(1/|x|)$. If $x = 0$, $1/x$ is not defined.

Now it is a routine, though tedious, matter to verify that with the above definitions of the arithmetic operations, all of the familiar laws governing these operations are valid. For example, to name a few:

1. For real numbers x_1 and x_2,
$$x_1 + x_2 = x_2 + x_1, \qquad \text{and} \qquad x_1x_2 = x_2x_1.$$
(The commutative laws for addition and multiplication, respectively.)
2. For real numbers x_1, x_2, and x_3,
$$(x_1 + x_2) + x_3 = x_1 + (x_2 + x_3), \quad \text{and} \quad (x_1x_2)x_3 = x_1(x_2x_3).$$
(The associative laws for addition and multiplication, respectively.)
3. For real numbers x_1, x_2, and x_3,
$$x_1(x_2 + x_3) = x_1x_2 + x_1x_3.$$
(The distributive law.)

Furthermore, it is a routine matter to verify that the usual laws governing the manipulation of inequalities, the absolute value function, etc., carry over

without change to the general context of the real number system. For example, the *triangle inequality* is true in the real number system, and states that for any real numbers x_1 and x_2, it is always true that $|x_1 + x_2| \leq |x_1| + |x_2|$. Because the proofs of these results are all very similar and follow in rather obvious ways from the corresponding facts for rational numbers, we omit them here.

There is, however, a very important property of the real number system that is not shared by the rational number system, and which is described in the following theorem.

Theorem 1.1 *Suppose S is a nonempty set of real numbers which is bounded above. That is, there exists a number M, such that $x \leq M$ for each x in S. Then there exists exactly one real number s such that*

1. *$s' \leq s$, for each s' in S.*
2. *No number less than s has this property. That is, for any number $s^* < s$, there exists a number s' in S such that $s^* < s'$.*

The number s is called the supremum, or least upper bound, of the set S.

PROOF Clearly, there can be at most one such number. For suppose s_1 and s_2 are distinct numbers having the properties described above. Then either $s_1 < s_2$ or $s_2 < s_1$. In either case, this would contradict the second requirement above.

In order to exhibit a number having the desired properties, we simply take the cut that is the set-theoretic union of the cuts in S.

Remark If a set of real numbers is not bounded above, we say that its supremum is ∞.

In many cases, Theorem 1.1 can make precise our notion of what it means for an infinite series to converge. For example, suppose $a_0, a_1, a_2, \ldots$ is a sequence of non-negative numbers about which it is in some way known that the associated sequence of "partial sums" $s_0, s_1, s_2, \ldots$ is bounded, where $s_0 = a_0$, $s_1 = a_0 + a_1$, $s_2 = a_0 + a_1 + a_2$, etc. Then Theorem 1.1 asserts that there exists a unique number s, which is the smallest number that is less than or equal to all the s_n's. Upon reflection, it is perfectly natural to define s to be the value of the infinite "sum" $a_0 + a_1 + a_2 + \cdots$, and this is, in fact, precisely the definition used in practice. As a practical matter, we

may ask how it could ever be known a priori that the sequence $s_0, s_1, s_2, \ldots$ is bounded above. This information, however, is frequently easy to obtain. For example, suppose r is a number such that $0 \leq r < 1$, and define $a_0 = 1$, $a_1 = r$, $a_2 = r^2$, and, in general, $a_n = r^n$. Then, from high-school algebra,

$$\begin{aligned} s_n &= 1 + r + \cdots + r^n \\ &= \frac{1 - r^{n+1}}{1 - r} \\ &= \frac{1}{1 - r} - \frac{r^{n+1}}{1 - r}. \end{aligned}$$

In particular, therefore, $s_n \leq 1/(1 - r)$ for all n, so in this case, it is clear that the s_n's are bounded above. Moreover, it is fairly obvious that the value of the infinite "geometric series" $1 + r + r^2 + \cdots$, in the sense of the previous remarks, is $1/(1 - r)$; we shall soon be in a position to verify this. However, before we pass on to questions of this sort, a Corollary to Theorem 1.1 bears mentioning.

Corollary to Theorem 1.1 *Suppose S is a nonempty set of real numbers that is bounded below. That is, suppose there exists a number M, such that $x \geq M$ for each x in S. Then there exists exactly one real number s such that*

1. *$s' \geq s$, for each s' in S.*
2. *No number greater than s has this property. That is, for any number $s^* > s$, there exists a number s' in S such that $s' < s^*$.*

The number s is called the infimum, or greatest lower bound, of the set S.

PROOF This result follows from Theorem 1.1, by replacing the set S in the Corollary by the set $-S$, whose elements consist of the negatives of the elements of S.

3 Infinite sequences and series

Suppose $s_0, s_1, s_2, \ldots$ is an infinite sequence of numbers. That is, suppose that for every non-negative integer n there is given some number s_n. Various notations are used to express this concept. One of them is to write explicitly, as we have done, a few terms of the sequence, followed by a series of dots.

Another is to write $\{s_n\}_{n=0}^{\infty}$ or simply $\{s_n\}$. Occasionally, it is convenient to take an integer k, other than 0, as the initial subscript, and to write, for example, $s_k, s_{k+1}, s_{k+2}, \ldots$ or $\{s_n\}_{n=k}^{\infty}$. In any event, the question we are interested in is the following: when does a sequence converge, or tend, to some definite number? To answer this, we must first give a precise meaning to the notion of convergence.

Definition 1.9 We say that a sequence $s_0, s_1, s_2, \ldots$ *converges* to a number s if and only if $|s - s_n|$ becomes arbitrarily small as n becomes arbitrarily large. More precisely, if and only if for every positive number ϵ, there exists an integer n_0, in general dependent on ϵ, such that for $n \geq n_0$, $|s - s_n| \leq \epsilon$. The number s is called the *limit* of the sequence. In graphical terms, the convergence of the sequence $s_0, s_1, s_2, \ldots$ to s means that the points corresponding to the s_n's cluster more and more closely about the point corresponding to s.

The statement that the sequence $s_0, s_1, s_2, \ldots$ converges to s is written in various ways, the most common ones being $s_n \to s$ and $\lim_{n\to\infty} s_n = s$.

It is obvious that $s_n \to s$ if and only if $s - s_n \to 0$, that is, if and only if the sequence whose nth term is $s - s_n$ converges to 0. It follows that $s_n \to s$ if and only if $s_n = s + \delta_n$, where $\delta_0, \delta_1, \delta_2, \ldots$ is a sequence that converges to 0.

Example

Suppose $|r| < 1$. Then $r^n \to 0$, or $s_n \to 0$, where $s_0, s_1, s_2, \ldots$ is the sequence whose nth term is r^n. (We adopt the convention that $0^0 = 1$, if $r = 0$.)

Proof: We may suppose that $r \neq 0$, since the claim is obvious in this case. Now it is easy to see that the statement is equivalent to the statement that $1/|r^n|$ becomes arbitrarily large as n becomes arbitrarily large. In other words, for any number $M > 0$, there exists a positive integer n_0 such that $1/|r^n| \geq M$ for all $n \geq n_0$. This is expressed by writing $1/|r^n| \to \infty$ as $n \to \infty$. In order to see that $1/|r^n| \to \infty$ as $n \to \infty$, we set $1/|r| = 1 + c$, for some $c > 0$. Then, by the binomial theorem,

$$\begin{aligned}\frac{1}{|r^n|} &= \left(\frac{1}{|r|}\right)^n \\ &= (1 + c)^n \\ &= 1 + \binom{n}{1}c + \cdots + \binom{n}{n-1}c^{n-1} + c^n.\end{aligned}$$

In particular,

$$\frac{1}{|r^n|} > 1 + \binom{n}{1}c = 1 + nc,$$

and the desired result follows immediately from this.

Theorem 1.2 *Suppose $s_0, s_1, s_2, \ldots$ is a convergent sequence with limit s, and suppose $s'_0, s'_1, s'_2, \ldots$ is a subsequence of $s_0, s_1, s_2, \ldots$. That is, suppose there exists a sequence of non-negative integers $n_0, n_1, n_2, \ldots$, satisfying $n_0 < n_1 < n_2 < \cdots$, and such that $s'_k = s_{n_k}$. Then the sequence $s'_0, s'_1, s'_2, \ldots$ is convergent and converges to s.*

PROOF This is obvious, since $|s - s_n| \leq \epsilon$ for sufficiently large n, implies that $|s - s'_n| \leq \epsilon$ for sufficiently large n.

Theorem 1.3 *Suppose $s_0, s_1, s_2, \ldots$ is a convergent sequence, and suppose $s'_0, s'_1, s'_2, \ldots$ is a subsequence that converges to s. Then $s_0, s_1, s_2, \ldots$ converges to s.*

PROOF This follows immediately from Theorem 1.2.

Theorem 1.4 *Suppose $\{s_n\}$ is a convergent sequence with limit s, and suppose c is a constant. Then the sequence $\{cs_n\}$ is convergent and converges to cs.*

PROOF If $c = 0$, the assertion is obvious. Accordingly, suppose $c \neq 0$. We must show that for any $\epsilon > 0$, there exists an integer n_0 such that for $n \geq n_0$, $|cs - cs_n| \leq \epsilon$. Now it follows from the fact that $s_n \to s$, that for any $\epsilon > 0$, there exists an integer n_0, such that $|s - s_n| \leq \epsilon/|c|$ for $n \geq n_0$. Thus, for $n \geq n_0$, $|cs - cs_n| = |c|\,|s - s_n| \leq \epsilon$, which proves the assertion.

Theorem 1.5 *Suppose $\{s_n\}$ and $\{s'_n\}$ are two convergent sequences, with limits s and s', respectively. Then the sequence $\{s_n + s'_n\}$ is convergent and converges to $s + s'$.*

PROOF We must show that for any $\epsilon > 0$, there exists an integer n_0 such that for $n \geq n_0$, $|(s + s') - (s_n + s'_n)| \leq \epsilon$. Now $(s + s') - (s_n + s'_n) = (s - s_n) + (s' - s'_n)$. Moreover, since $s_n \to s$ and $s'_n \to s'$, it follows that for any $\epsilon > 0$, there exist integers n_1 and n_2 such that for $n \geq n_1$, $|s - s_n| \leq \frac{1}{2}\epsilon$, and for $n \geq n_2$, $|s' - s'_n| \leq \frac{1}{2}\epsilon$. Thus, if we define $n_0 = \max(n_1, n_2)$, it is

clear, since $|(s - s_n) + (s' - s'_n)| \leq |s - s_n| + |s' - s'_n|$, that for $n \geq n_0$, $|(s + s') - (s_n + s'_n)| \leq \epsilon$, which proves the theorem.

Corollary *Suppose $\{s_n\}$ and $\{s'_n\}$ are two convergent sequences with limits s and s', respectively. Then the sequence $\{s_n - s'_n\}$ is convergent and converges to $s - s'$.*

PROOF This follows immediately from the last two theorems.

Theorem 1.6 *Suppose $\{s_n\}$ and $\{s'_n\}$ are two convergent sequences with limits s and s', respectively. Then the sequence $\{s_n s'_n\}$ is convergent and converges to ss'.*

PROOF By hypothesis, there exist sequences $\{\delta_n\}$ and $\{\delta'_n\}$, both converging to 0, such that $s_n = s + \delta_n$ and $s'_n = s' + \delta'_n$. Thus, $s_n s'_n = (s + \delta_n)(s' + \delta'_n) = ss' + s'\delta_n + s\delta'_n + \delta_n\delta'_n$. Now the sequences $\{s'\delta_n\}$, $\{s\delta'_n\}$, and $\{\delta_n\delta'_n\}$ all converge to 0; the first two by Theorem 1.4 and the last because $|\delta_n\delta'_n| \leq |\delta_n|$ for n beyond some point. Thus, since Theorem 1.5 can obviously be extended to any finite number of sums, and in particular, three, it follows that $ss' + s'\delta_n + s\delta'_n + \delta_n\delta'_n \to ss'$.

Theorem 1.7 *Suppose $\{s_n\}$ is a convergent sequence with limit s, where $s \neq 0$. Suppose, moreover, that none of the s_n's is zero. Then the sequence $\{1/s_n\}$ is convergent, and converges to $1/s$.*

PROOF By hypothesis, $s_n = s + \delta_n$, with $\delta_n \to 0$. Now

$$\frac{1}{s} - \frac{1}{s + \delta_n} = \frac{\delta_n}{s(s + \delta_n)}.$$

Moreover, since $\delta_n \to 0$ and $s \neq 0$, it is evident that for n beyond some point, $|s + \delta_n| \geq |s/2|$, so for sufficiently large n, $|\delta_n/s(s + \delta_n)| \leq 2|\delta_n|/s^2$. Since $2|\delta_n|/s^2 \to 0$, this immediately implies the desired result.

Suppose now that $a_0, a_1, a_2, \ldots$ is a sequence of numbers. We can, as before, associate with the original sequence, the sequence $s_0, s_1, s_2, \ldots$ of *partial sums*, defined by $s_0 = a_0$, $s_1 = a_0 + a_1$, $s_2 = a_0 + a_1 + a_2$, etc. Thus, in the customary notation, $s_n = \sum_{k=0}^{n} a_k$.

Now one can naturally ask what meaning, if any, can be assigned to the formal expression $\sum_{n=0}^{\infty} a_n$. Such an expression is called an *infinite series*, and the a_n's are called the *terms* of the series.

Definition 1.10 An infinite series $\sum_{n=0}^{\infty} a_n$ is said to *converge* to the value s, if and only if the associated sequence $\{s_n\}$ of partial sums converges to s. The number s is called the *sum*, or *limit*, of the series. If a series does not converge to any number, it is said to *diverge*.

Theorem 1.8 *Suppose $\sum_{n=0}^{\infty} a_n$ is a series whose terms are non-negative, that is, such that $a_n \geq 0$ for each n. Then if the associated sequence $\{s_n\}$ of partial sums is bounded above, that is, if there exists a number M such that $s_n \leq M$ for all n, then the series $\sum_{n=0}^{\infty} a_n$ is convergent and converges to s, the supremum of the set of numbers $\{s_0, s_1, s_2, \ldots\}$. If, on the other hand, the sequence $\{s_n\}$ is not bounded above, the series $\sum_{n=0}^{\infty} a_n$ diverges.*

PROOF If the partial sums are not bounded above, it is obvious that the series diverges. Accordingly, suppose the partial sums are bounded above, with supremum s. Now since the a_n's are non-negative, it is clear that the partial sums are increasing, that is, $s_0 < s_1 < s_2 \leq \cdots$. Moreover, $s_n \leq s$ for each n. This implies that in order to prove the theorem, it suffices to show that for any $\epsilon > 0$, there exists an integer n_0 such that $s - s_{n_0} \leq \epsilon$, since the inequality must continue to be true if n_0 is replaced by any $n \geq n_0$. But for a given ϵ, the existence of such an n_0 is obvious; for, if there were no n_0 such that $s - s_{n_0} \leq \epsilon$, then s could clearly not be the supremum of the set of numbers $\{s_0, s_1, s_2, \ldots\}$, that is, the *least* number that stands in the relation $\geq$ to all the s_n's.

Remark We have already seen that if $0 \leq r < 1$, the terms of the geometric series $\sum_{n=0}^{\infty} r^n$ are bounded above, so the last theorem implies that this series converges. Moreover, since

$$1 + r + \cdots + r^n = \frac{1}{1-r} - \frac{r^{n+1}}{1-r}$$

and $r^{n+1}/(1-r) \to 0$, it is clear that the series converges to $1/(1-r)$.

There are various tests for convergence based on Theorem 1.8. For the sake of initial simplicity, we shall temporarily confine our attention to the case of infinite series with non-negative terms.

Theorem 1.9 (The Majorization, or Comparison Test) *Suppose $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ are series with non-negative terms. Suppose, moreover, that $\sum_{n=0}^{\infty} a_n$ converges and that $\sum_{n=0}^{\infty} b_n$ is majorized by $\sum_{n=0}^{\infty} a_n$, that is, $b_n \leq a_n$ for each n. Then $\sum_{n=0}^{\infty} b_n$ converges.*

PROOF The proof follows immediately from Theorem 1.8. Since $\sum_{n=0}^{\infty} a_n$ is convergent, its partial sums are bounded above, and since $b_n \le a_n$ for each n, this implies that the partial sums of $\sum_{n=0}^{\infty} b_n$ are bounded above, so the last series must converge.

Theorem 1.10 (The Ratio Test) *Suppose $\sum_{n=0}^{\infty} a_n$ is a series whose terms beyond some point are all positive. Suppose, moreover, that there exists some number r such that $0 < r < 1$ and such that for all n beyond some point $a_{n+1}/a_n \le r$. Then the series $\sum_{n=0}^{\infty} a_n$ is convergent.*

PROOF Since the values of a finite number of terms cannot possibly affect the convergence of a series, we may assume, by discarding the first few terms if necessary, that $a_{n+1}/a_n \le r$ for all n. Then $a_1 \le a_0 r$, $a_2 \le a_1 r$, $a_3 \le a_2 r$, etc., which implies that $a_n \le a_0 r^n$ for all n. This implies that $\sum_{n=0}^{\infty} a_n$ is majorized by the convergent geometric series $\sum_{n=0}^{\infty} a_0 r^n$, which proves the result.

Example

The series

$$\sum_{n=0}^{\infty} \frac{1}{n!} = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots$$

converges by the ratio test.

Remark It may sometimes happen that the ratios a_{n+1}/a_n are all less than 1, but there is no number of the type described in Theorem 1.10. For example, for the series

$$\sum_{n=1}^{\infty} \frac{1}{n}, \qquad \frac{a_{n+1}}{a_n} = \frac{n}{n+1} = \frac{1}{1 + (1/n)}$$

which tends to 1, so such an r cannot exist. As we shall shortly see, the series $\sum_{n=1}^{\infty} 1/n$ is, in fact, divergent. On the other hand, the series $\sum_{n=1}^{\infty} 1/n^2$, for which the ratios again tend to 1, is convergent. We conclude that the mere information that $a_{n+1}/a_n < 1$ for all n, is, by itself, insufficient to infer anything about the convergence or divergence of the series in question.

Theorem 1.11 (The Root Test) *Suppose $\sum_{n=0}^{\infty} a_n$ is a series whose terms beyond some point are non-negative. Suppose, moreover, that there exists a number r such that $0 < r < 1$ and such that for all n beyond some point, $a_n^{1/n} \le r$. Then the series $\sum_{n=0}^{\infty} a_n$ is convergent.*

PROOF The proof follows immediately from the majorization test, since it is clear that from some point onwards, the terms of the series $\sum_{n=0}^{\infty} a_n$ are dominated by those of the convergent series $\sum_{n=0}^{\infty} r^n$.

Example

The series $\sum_{n=0}^{\infty} 1/n^n$ converges by the root test, as does the series

$$\sum_{n=0}^{\infty} \frac{1}{[1 + (1/n)]^{n^2}}.$$

(In the latter case, we assume as known the fact that $[1 + (1/n)]^n \to e \approx 2.718 \cdots$.)

Remark As in the case of the ratio test, the mere information that $a_n^{1/n} < 1$ from some point onwards does not imply anything about the convergence or divergence of a series. In fact, the series $\sum_{n=1}^{\infty} 1/n$ and $\sum_{n=1}^{\infty} 1/n^2$ again serve to illustrate this point.

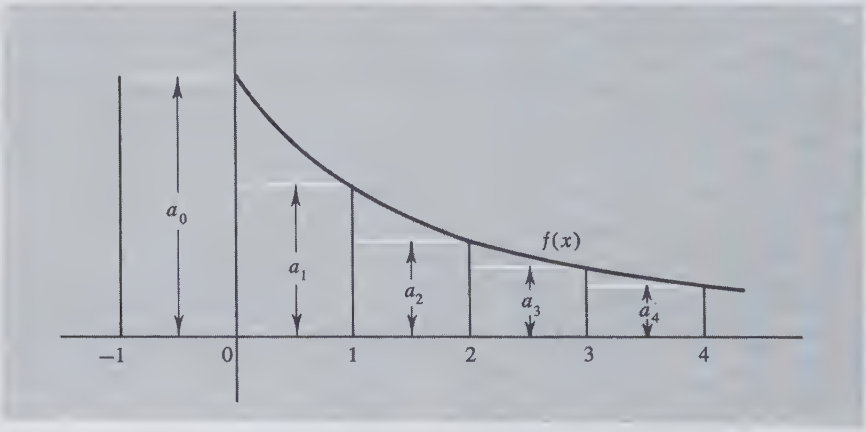

FIGURE 1

Theorem 1.12 (The Integral Test) [For this test, we assume familiarity with the notion of an improper integral.] *Suppose $\sum_{n=0}^{\infty} a_n$ is a series whose terms are non-negative and decreasing, that is, $a_0 \geq a_1 \geq a_2 \geq \cdots$. Suppose, moreover, that $f(x)$ is a non-negative decreasing function defined for $x \geq 0$ and such that $f(n) = a_n$. Then, if the integral $\int_0^{\infty} f(x)\,dx$ is finite, $\sum_{n=0}^{\infty} a_n$ is convergent. On the other hand, if $\int_0^{M} f(x)\,dx \to \infty$ as $M \to \infty$, the series $\sum_{n=0}^{\infty} a_n$ diverges.*

PROOF Suppose $\int_0^{\infty} f(x)\,dx$ is finite. Now it is evident, from Figure 1, that the partial sums of $\sum_{n=0}^{\infty} a_n$ are equal in numerical value to the aggregate

area of the corresponding set of blocks. The aggregate area of any number of these blocks is clearly less than or equal to $a_0 + \int_0^\infty f(x)\,dx$, so the partial sums of $\sum_{n=0}^\infty a_n$ are bounded above, and hence, the series converges. If, on the other hand $\int_0^\infty f(x)\,dx$ is infinite, it is evident from an inspection of Figure 2 that the partial sums of $\sum_{n=0}^\infty a_n$ must be unbounded, and hence, the series must diverge.

Example

For $p > 1$, the integral $\int_0^\infty 1/(1 + x)^p\,dx$ is finite. Thus, for $p > 1$, the series $\sum_{n=0}^\infty 1/(1 + n)^p$, or, what is the same thing, $\sum_{n=1}^\infty 1/n^p$ converges. On the other hand, if $p = 1$, the integral

$$\int_0^\infty \frac{1}{1 + x}\,dx = \lim_{M\to\infty} \int_0^M \frac{1}{1 + x}\,dx = \lim_{M\to\infty} [\log (M + 1)]$$

is infinite, so $\sum_{n=1}^\infty 1/n$ diverges.

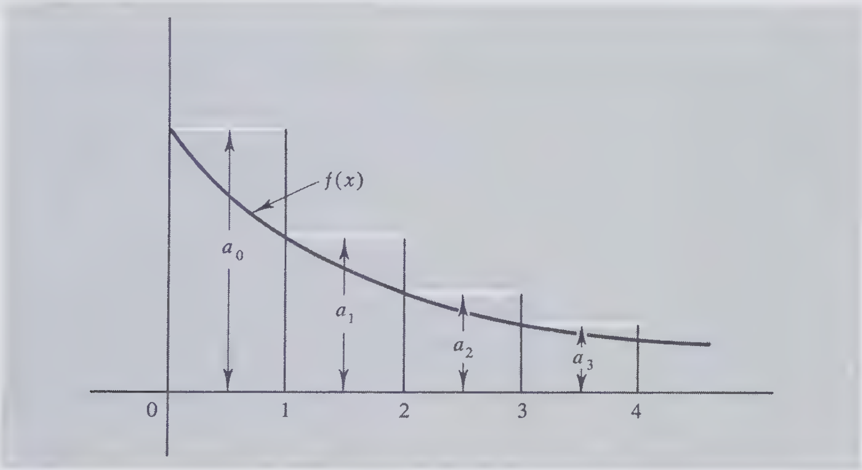

FIGURE 2

Up to this point, we have restricted our attention to series whose terms are non-negative. The basic advantage of such series lies in the fact that if one can, in one way or another, determine that the partial sums are bounded, then the series must converge. Now if no restrictions are placed on the signature of the terms of an infinite series, this may no longer be true. For example, the partial sums of the series $\sum_{n=0}^\infty (-1)^n = 1 - 1 + 1 - 1 + \cdots$ are bounded above (and also below), but the series does not converge. There are, nevertheless, several useful criteria for determining whether or not an infinite series converges, regardless of restrictions on the signature of its terms.

Theorem 1.13 *Suppose $\sum_{n=0}^{\infty} a_n$ is an infinite series, and suppose $\sum_{n=0}^{\infty} |a_n|$ converges. Then $\sum_{n=0}^{\infty} a_n$ converges.*

PROOF Define numbers $b_0, b_1, b_2, \ldots$ by setting $b_n = |a_n| + a_n$. Then $b_n \geq 0$ for $n = 0, 1, 2, \ldots$. Moreover, the series $\sum_{n=0}^{\infty} b_n$ converges, since its partial sums are obviously dominated by those of the convergent series $\sum_{n=0}^{\infty} 2|a_n|$. Thus, since $\sum_{n=0}^{\infty} a_n$ is expressible as the difference of two convergent series, namely $\sum_{n=0}^{\infty} b_n$ and $\sum_{n=0}^{\infty} |a_n|$, it follows from the Corollary to Theorem 1.5 that $\sum_{n=0}^{\infty} a_n$ is convergent. [The particular version of the Corollary that we are using here is explicitly stated as Theorem 1.18 in Section 4.]

Definition 1.11 Suppose $\sum_{n=0}^{\infty} a_n$ is an infinite series. If the series $\sum_{n=0}^{\infty} |a_n|$ converges, then $\sum_{n=0}^{\infty} a_n$ is said to be *absolutely convergent.* A series that is convergent, but not absolutely convergent, is called *conditionally convergent.*

Remark As Theorem 1.13 shows, absolute convergence implies convergence. The reverse is not true, however. A series may be convergent without being absolutely convergent. An example is given by the series $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots$, which is convergent, but not absolutely convergent. The fact that this series is convergent is a consequence of the following theorem.

Theorem 1.14 (Alternating Series Test) *Suppose $a_0, a_1, a_2, \ldots$ is a sequence such that $a_n \to 0$, and the a_n's are decreasing, that is, $a_0 \geq a_1 \geq a_2 \geq \cdots$. Then the series $a_0 - a_1 + a_2 - a_3 + a_4 - a_5 + \cdots = \sum_{n=0}^{\infty} (-1)^n a_n$ is convergent.*

PROOF Denote by $s_0, s_1, s_2, \ldots$ the sequence of partial sums associated with the series. Then, if n is odd,

$$\begin{aligned} s_n &= (a_0 - a_1) + (a_2 - a_1) + \cdots + (a_{n-1} - a_n) \\ &= a_0 - (a_1 - a_2) - \cdots - (a_{n-2} - a_{n-1}) - a_n. \end{aligned}$$

Taken together, these last two equalities show that the sequence $s_1, s_3, s_5, \ldots = \{s_{2n+1}\}$ of odd partial sums is increasing and bounded above, and hence convergent. Thus, since $s_{2n} = s_{2n+1} + a_{2n+1}$, and since $a_{2n+1} \to 0$, it follows that the sequence $s_0, s_2, s_4, \ldots$ of even partial sums is also convergent, and converges to the same limit. This clearly implies that the sequence $s_0, s_1, s_2, s_3, s_4, \ldots$ of all partial sums is convergent, and converges to the joint limit of the odd and even partial sums.

The following simple result, while by no means a criterion for convergence, is sometimes helpful.

Theorem 1.15 *Suppose $\sum_{n=0}^{\infty} a_n$ is a convergent series. Then it must be the case that $a_n \to 0$.*

PROOF Denote by $s_0, s_1, s_2, \ldots$ the partial sums of the series $\sum_{n=0}^{\infty} a_n$. Then, since $\sum_{n=0}^{\infty} a_n$ is convergent, the sequences $s_0, s_1, s_2, \ldots$ and $0, s_0, s_1, s_2, \ldots$ must both be convergent, and converge to the same limit. Now for each n, a_n is the difference between the nth elements of the two sequences, and by the Corollary to Theorem 1.5, this implies the desired result.

Remark It may very well happen that a series $\sum_{n=0}^{\infty} a_n$ diverges, and at the same time $a_n \to 0$. For example, the series $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots$ has this property. Theorem 1.15 does not preclude this. It simply states that if a series is convergent to begin with, then its terms must tend to zero.

4 Algebraic operations with series

Theorem 1.16 *Suppose $\sum_{n=0}^{\infty} a_n$ is a convergent series with limit s. Then for any constant c, the series $\sum_{n=0}^{\infty} ca_n$ is convergent and converges to cs.*

PROOF This follows immediately from Theorem 1.4.

Theorem 1.17 *Suppose $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ are convergent series with limits s and s', respectively. Then the series $\sum_{n=0}^{\infty} (a_n + b_n)$ is convergent and converges to $s + s'$.*

PROOF This follows immediately from Theorem 1.5.

Theorem 1.18 *Suppose $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ are convergent series with limits s and s', respectively. Then the series $\sum_{n=0}^{\infty} (a_n - b_n)$ is convergent and converges to $s - s$.*

PROOF This follows immediately from the Corollary to Theorem 1.5.

The question of products is somewhat more complicated. As a preliminary result, we require the following theorem, which is interesting in its own right.

Theorem 1.19 *Suppose $\sum_{n=0}^{\infty} a_n$ is an absolutely convergent series, and suppose $\sum_{n=0}^{\infty} b_n$ is a rearrangement of $\sum_{n=0}^{\infty} a_n$, by which we simply mean that there is a 1–1 correspondence between the terms of the two series. In other words, there exists a sequence $n_0, n_1, n_2, \ldots$ of non-negative integers, in which each non-negative integer appears exactly once, such that $b_0 = a_{n_0}$, $b_1 = a_{n_1}$, $b_2 = a_{n_2}, \ldots$. Then $\sum_{n=0}^{\infty} b_n$ is absolutely convergent, and has the same sum as $\sum_{n=0}^{\infty} a_n$.*

Remark The requirement that $\sum_{n=0}^{\infty} a_n$ be *absolutely convergent* is essential (*cf.* problem 11).

PROOF To begin with, suppose that the terms of $\sum_{n=0}^{\infty} a_n$ are non-negative. Then the sum of $\sum_{n=0}^{\infty} a_n$, which we denote by s, is simply the supremum of the partial sums $s_0, s_1, s_2, \ldots$. Since the terms of $\sum_{n=0}^{\infty} b_n$ are also non-negative, it is sufficient in this case to show that the supremum of the partial sums of $\sum_{n=0}^{\infty} b_n$ is finite and equals s. To show that this must be so, suppose that the supremum of the partial sums $s'_0, s'_1, s'_2, \ldots$ of $\sum_{n=0}^{\infty} b_n$ is greater than s. Then there exists some m such that $s'_m > s$. On the other hand, since $\sum_{n=0}^{\infty} b_n$ is a rearrangement of $\sum_{n=0}^{\infty} a_n$, it is clear that there exists some n such that $s_n \geq s'_m$. This is impossible, however, since $s_n \leq s$.

Suppose now that the supremum of $s'_0, s'_1, s'_2, \ldots$ is less than s. Then there exists an $\epsilon > 0$ such that $s'_n \leq s - \epsilon$ for all n. Now since s is the supremum of $s_0, s_1, s_2, \ldots$, there exists some m such that $s_m > s - \epsilon$. On the other hand, since $\sum_{n=0}^{\infty} b_n$ is a rearrangement of $\sum_{n=0}^{\infty} a_n$, it is clear that there exists some n such that $s'_n \geq s_m$, which is impossible, since $s'_n \leq s - \epsilon$. We conclude from this contradiction that the supremum of $s'_0, s'_1, s'_2, \ldots$ must be s, and hence the series $\sum_{n=0}^{\infty} b_n$ converges to s.

To establish the theorem for the general case where the a_n's can be of any signature, note first of all that the above reasoning shows that $\sum_{m=0}^{\infty} b_m$ must be absolutely convergent, and hence convergent. Next observe that $\sum_{n=0}^{\infty} a_n$ can be written as the difference of two series $\sum_{n=0}^{\infty} a'_n$ and $\sum_{n=0}^{\infty} a''_n$, where $a'_n = a_n$ if $a_n \geq 0$, and $a'_n = 0$ if $a_n < 0$, while $a''_n = -a_n$ if $a_n < 0$, and $a''_n = 0$ if $a_n \geq 0$. Similarly, $\sum_{n=0}^{\infty} b_n$ can be expressed in the above way as the difference of two series $\sum_{n=0}^{\infty} b'_n$ and $\sum_{n=0}^{\infty} b''_n$, and it is clear that $\sum_{n=0}^{\infty} b'_n$ is a rearrangement of $\sum_{n=0}^{\infty} a'_n$, and $\sum_{n=0}^{\infty} b''_n$ is a rearrangement of $\sum_{n=0}^{\infty} a''_n$. By what we have already proved, $\sum_{n=0}^{\infty} a'_n = \sum_{n=0}^{\infty} b'_n$ and $\sum_{n=0}^{\infty} a''_n = \sum_{n=0}^{\infty} b''_n$. This immediately establishes the theorem, since $\sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty} a'_n - \sum_{n=0}^{\infty} a''_n$ and $\sum_{n=0}^{\infty} b_n = \sum_{n=0}^{\infty} b'_n - \sum_{n=0}^{\infty} b''_n$.

We now take up the question of products. To begin with, suppose that $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ are two absolutely convergent series with sums s and s',

respectively. If we formally multiply the series, we obtain an infinite collection of summands, which can be conveniently displayed as follows

$$
\begin{matrix}
a_0b_0 & a_0b_1 & a_0b_2 & a_0b_3 & \cdots \\
a_1b_0 & a_1b_1 & a_1b_2 & a_1b_3 & \cdots \\
a_2b_0 & a_2b_1 & a_2b_2 & a_2b_3 & \cdots \\
a_3b_0 & a_3b_1 & a_3b_2 & a_3b_3 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \cdots
\end{matrix}
\tag{1.1}
$$

Next we select a method of enumerating this array. That is, of setting up a 1–1 correspondence between the integers 1, 2, 3, . . . and the elements of the array. There are many ways in which this can be accomplished. One method, in which the elements are counted in the order indicated by the arrows, is illustrated below.

$$
\begin{tikzcd}
a_0b_0 \arrow[r] & a_0b_1 \arrow[dl] & a_0b_2 \arrow[r] & a_0b_3 \arrow[dl] & \cdots \\
a_1b_0 \arrow[d] & a_1b_1 \arrow[ur] & a_1b_2 \arrow[dl] & \cdots & \cdots \\
a_2b_0 \arrow[ur] & a_2b_1 \arrow[dl] & \cdots & \cdots & \cdots \\
a_3b_0 \arrow[d] & \cdots & \cdots & \cdots & \cdots \\
\vdots & \vdots & \vdots & \vdots & \cdots
\end{tikzcd}
$$

Having chosen some method of enumeration, we let c_n be the element of the array corresponding to the integer n. We shall show that the series $\sum_{n=0}^{\infty} c_n$ is absolutely convergent, and converges to ss'. We begin by showing that the series must be absolutely convergent; to do this, it suffices to show that the partial sums $s_0, s_1, s_2, \ldots$ of $\sum_{n=0}^{\infty} |c_n|$ must be bounded above. Now a given partial sum s_n consists of a finite sum of terms of the form $|a_jb_k|$. For a given s_n, denote by N the largest of the j's and k's which appear as subscripts in this way. Then clearly

$$s_n \le \left(\sum_{n=0}^{N} |a_n|\right)\left(\sum_{n=0}^{N} |b_n|\right).$$

Now by hypothesis, both $\sum_{n=0}^{\infty} |a_n|$ and $\sum_{n=0}^{\infty} |b_n|$ converge, and if we denote their sums by S and S', respectively, it follows that $s_n \leq SS'$ for all n, which shows that $\sum_{n=0}^{\infty} c_n$ is absolutely convergent.

It remains to be shown that $\sum_{n=0}^{\infty} c_n$ converges to ss'. To show this, we consider the sequence $d_0, d_1, d_2, \ldots$, which is obtained by enumerating (1.1) in the following way:

$$\begin{tikzcd}
a_0b_0 \arrow[r] & a_0b_1 \arrow[d] & a_0b_2 \arrow[r] & a_0b_3 \arrow[d] & \cdots \\
a_1b_0 \arrow[d] & a_1b_1 \arrow[l] & a_1b_2 \arrow[u] & a_1b_3 \arrow[d] & \cdots \\
a_2b_0 \arrow[r] & a_2b_1 \arrow[r] & a_2b_2 \arrow[u] & a_2b_3 \arrow[d] & \cdots \\
a_3b_0 \arrow[d] & a_3b_1 \arrow[l] & a_3b_2 \arrow[l] & a_3b_3 \arrow[l] & \cdots \\
\vdots & \vdots & \vdots & \vdots & \cdots
\end{tikzcd}$$

By the preceding reasoning, $\sum_{n=0}^{\infty} d_n$ must be absolutely convergent. Moreover, it is evident that for $k \geq 1$, the $(k^2 - 1)$th partial sum of $\sum_{n=0}^{\infty} d_n$ is simply $(\sum_{n=0}^{k-1} a_n)(\sum_{n=0}^{k-1} b_n)$, which converges to ss'. Thus, $\sum_{n=0}^{\infty} d_n$ is a convergent series for which a subsequence of the sequence of partial sums converges to ss', and by Theorem 1.3, this implies that $\sum_{n=0}^{\infty} d_n$ itself converges to ss'. Since $\sum_{n=0}^{\infty} c_n$ is a rearrangement of $\sum_{n=0}^{\infty} d_n$, it follows that $\sum_{n=0}^{\infty} c_n$ converges to ss', which was to be proved.

Summarizing, we have shown that the sum of the terms that naturally occur in the product of two series will, when taken in any order whatsoever, give the product of the sums of the series, provided that the latter are both absolutely convergent. A consequence of this fact, which is, as we shall see, important in the study of power series, is that the so-called "Cauchy product" of two absolutely convergent series converges to the product of the sums of the factors. The terms of the Cauchy product correspond to the sums of the diagonals as illustrated.

$$\begin{array}{ccccccccc}
a_0b_0 & + & a_0b_1 & + & a_0b_2 & + & a_0b_3 & & \cdots \\
a_1b_0 & & a_1b_1 & & a_1b_2 & & \cdots & & \cdots \\
a_2b_0 & & a_2b_1 & & \cdots & & \cdots & & \cdots \\
a_3b_0 & & \cdots & & \cdots & & \cdots & & \cdots \\
\vdots & & \vdots & & \vdots & & \vdots & & \cdots
\end{array}$$

More precisely, given two series $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$, their Cauchy product is defined to be the series $\sum_{n=0}^{\infty} c_n$, where

$$c_n = a_0b_n + a_1b_{n-1} + \cdots + a_{n-1}b_1 + a_nb_0 = \sum_{k=0}^{n} a_kb_{n-k}.$$

Since the partial sums of the Cauchy product obviously constitute a subsequence of the partial sums of a series corresponding to an enumeration of (1.1), it follows from Theorem 1.2 that the Cauchy product must converge to the product of the sums of $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$, provided the latter are absolutely convergent. For reference, we state this as a theorem.

Theorem 1.20 *Suppose $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ are two absolutely convergent series having sums s and s', respectively. Then the series $\sum_{n=0}^{\infty} c_n$, where $c_n = \sum_{n=0}^{\infty} a_kb_{n-k}$, converges to ss'.*

Remark A stronger result than Theorem 1.20 is true. Namely, even if only one of the series $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$ is absolutely convergent, the Cauchy product of the two still converges to the product of the respective sums, provided both series are convergent (*cf.* [2], p. 321). On the other hand, if both series are merely conditionally convergent, it may happen that the Cauchy product diverges (*cf.* problem 15). Finally, it can be shown that if two series and their Cauchy product all converge, then the sum of the Cauchy product must be the product of the sums of the original two series (*cf.* [2], p. 321).

5 Fubini's theorem

The result of this section, which will be useful in Chapter 3, is a case of what has come to be known as Fubini's theorem. In the form in which we are presenting it, it is also known as Cauchy's double series theorem.

Suppose we are given an array of the form

$$\begin{array}{ccccc} a_{00} & a_{01} & a_{02} & a_{03} & \cdots \\ a_{10} & a_{11} & a_{12} & a_{13} & \cdots \\ a_{20} & a_{21} & a_{22} & a_{23} & \cdots \\ a_{30} & a_{31} & a_{32} & a_{33} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \cdots \end{array} \tag{1.2}$$

and suppose, moreover, that the series corresponding to each row is absolutely convergent. That is, $\sum_{n=0}^{\infty} |a_{mn}|$ is convergent for $m = 0, 1, 2, \ldots$. Set $r_m = \sum_{n=0}^{\infty} |a_{mn}|$, and assume that $\sum_{m=0}^{\infty} r_m$ converges. Then it is obvious that $\sum_{n=0}^{\infty} a_{mn}$ converges for $m = 0, 1, 2, \ldots$, and that if we set $r_m^* = \sum_{n=0}^{\infty} a_{mn}$, then $\sum_{m=0}^{\infty} r_m^*$ converges. Denote by L the sum of the series $\sum_{m=0}^{\infty} r_m^*$.

Theorem 1.21 *Under the above conditions, the series corresponding to each column of* (1.2) *is absolutely convergent. That is,* $\sum_{m=0}^{\infty} |a_{mn}|$ *converges for* $n = 0, 1, 2, \ldots$. *Moreover, if we set* $s_n = \sum_{m=0}^{\infty} |a_{mn}|$, *then* $\sum_{n=0}^{\infty} s_n$ *is convergent. Finally, and most importantly, if we set* $s_n^* = \sum_{m=0}^{\infty} a_{mn}$, *then* $\sum_{n=0}^{\infty} s_n^*$ *converges to L. That is,* $\sum_{m=0}^{\infty} (\sum_{n=0}^{\infty} a_{mn}) = \sum_{n=0}^{\infty} (\sum_{m=0}^{\infty} a_{mn})$.

PROOF We first prove the theorem for the case in which all the a_{mn}'s are non-negative. Then it is clear that $a_{mn} \le r_m$, for any m and n, and indeed, $a_{m0} + a_{m1} + \cdots + a_{mn} \le r_m$. In particular, therefore, the series $\sum_{m=0}^{\infty} a_{mn}$ corresponding to each column converges by the majorization test, since $\sum_{m=0}^{\infty} r_m$ converges. Moreover, for any n, $s_0^* + s_1^* + \cdots + s_n^* \le \sum_{m=0}^{\infty} r_m = L$. Since the s_n^*'s are non-negative, this implies that $\sum_{n=0}^{\infty} s_n^*$ converges and that the sum is $\le L$. Denote the sum by L'. Then a reversal of the preceding argument shows that $L \le L'$, so we conclude that $L = L'$, which proves the theorem if the a_{mn}'s are non-negative.

It remains to establish the theorem for the case where the a_{mn}'s are not necessarily non-negative. Note first that by the above reasoning, $\sum_{n=0}^{\infty} (\sum_{m=0}^{\infty} |a_{mn}|)$ and hence $\sum_{n=0}^{\infty} (\sum_{m=0}^{\infty} a_{mn})$ are both convergent. That is, in each case, both the inner series and the outer series are convergent. Next note that the array $\{a_{mn}\}$ can be expressed as the "difference" of two arrays $\{a'_{mn}\}$ and $\{a''_{mn}\}$, both of which have non-negative terms, and both of which satisfy the hypotheses of Theorem 1.21. (By the statement that $\{a_{mn}\}$ is the "difference" of $\{a'_{mn}\}$ and $\{a''_{mn}\}$, we simply mean that $a_{mn} = a'_{mn} - a''_{mn}$, for every possible combination of m and n.) To see that $\{a_{mn}\}$ can be expressed in this way, we simply define a'_{mn} to be a_{mn} if $a_{mn} \ge 0$, and 0 otherwise, and define a''_{mn} to be $-a_{mn}$ if $a_{mn} < 0$, and 0 otherwise.

Now

$$\begin{aligned}
\sum_{m=0}^{\infty}\left(\sum_{n=0}^{\infty} a_{mn}\right) &= \sum_{m=0}^{\infty}\left(\sum_{n=0}^{\infty} (a'_{mn} - a''_{mn})\right) \\
&= \sum_{m=0}^{\infty}\left(\sum_{n=0}^{\infty} a'_{mn}\right) - \sum_{m=0}^{\infty}\left(\sum_{n=0}^{\infty} a''_{mn}\right) \\
&= \sum_{n=0}^{\infty}\left(\sum_{m=0}^{\infty} a'_{mn}\right) - \sum_{n=0}^{\infty}\left(\sum_{m=0}^{\infty} a''_{mn}\right)
\end{aligned}$$

$$= \sum_{n=0}^{\infty} \left(\sum_{m=0}^{\infty} (a'_{mn} - a''_{mn}) \right)$$

$$= \sum_{n=0}^{\infty} \left(\sum_{m=0}^{\infty} a_{mn} \right),$$

which proves the theorem.

Exercises

1 Suppose $x \neq 0$ is rational and y is irrational. Show that $x + y$ and xy must be irrational.

2 Show that for any positive integer n that is not a perfect square, $\sqrt{n}$ is irrational.

3 Determine whether or not the following series converge.

a. $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n \log n}$ b. $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n(\log n)^2}$ c. $\displaystyle\sum_{n=1}^{\infty} \left(\frac{n}{n+1}\right)^{n^2}$

d. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^{1+1/n}}$ e. $\displaystyle\sum_{n=1}^{\infty} \sin n$

4 For any number x, show that the sequence $x_1, x_2, x_3, \ldots$ is convergent, where $x_1 = x$, $x_2 = \cos x_1$, $x_3 = \cos x_2$, and in general, $x_{n+1} = \cos x_n$. (*Hint*: $x_n = x_1 + (x_2 - x_1) + \cdots + (x_n - x_{n-1})$. Now $x_{k+1} - x_k = \cos x_k - \cos x_{k-1} = -(x_k - x_{k-1}) \sin y$, by the mean value theorem, where y is some number between x_{k-1} and x_k.)

5 Evaluate the sum of the series $\sum_{n=1}^{\infty} 1/n(n+1)$, using the fact that $1/n(n+1) = 1/n - 1/(n+1)$.

6 Suppose $s_1, s_2, s_3, \ldots$ is a convergent sequence with limit s. Show that the sequence $s_1^*, s_2^*, s_3^*, \ldots$, where $s_n^* = (s_1 + \cdots + s_n)/n$, is also convergent, and converges to s. Is the converse true? That is, if the sequence $s_1^*, s_2^*, s_3^*, \ldots$, derived from $s_1, s_2, s_3, \ldots$ in the above way, is convergent and converges to s, does it follow that $s_1, s_2, s_3, \ldots$ is convergent, and converges to s?

7 It is a theorem in plane geometry that the medians of a triangle intersect at a point. One somewhat unorthodox way of proving this is suggested by Figure 3. Since the diagonals of a parallelogram bisect each other, it is evident that the medians of any one of the central triangles coincide, where jointly defined, with the medians of the original triangle, and since the diameters of the central triangles tend to zero, this proves that the medians intersect at a point. Prove, by constructing an appropriate series, that the distance of the point of intersection from a vertex of the original triangle is two-thirds of the length of the median that connects them.

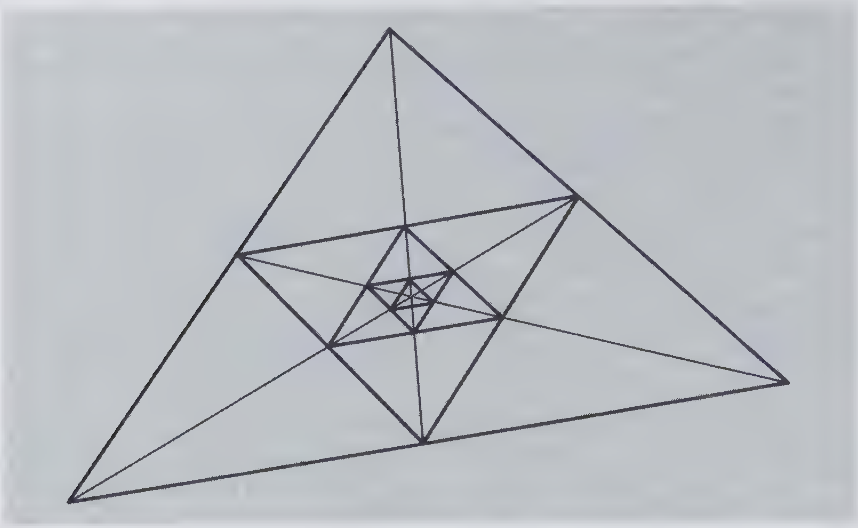

FIGURE 3

8 For $x > 0$, and a positive integer n, define $x^{1/n}$ to be the cut whose upper section consists of all positive rationals r, satisfying $r^n > x$. Show that $n^{1/n} \to 1$. (*Hint*: for each positive integer n, $n^{1/n} = 1 + c_n$, for some $c_n > 0$. Thus $(1 + c_n)^n = n$, or

$$1 + \binom{n}{1}c_n + \cdots + \binom{n}{n-1}c_n^{n-1} + c_n^n = n.$$

Show that this implies that $c_n \to 0$.)

9 Using the result of problem 8, show that for any $r > 0$, $r^{1/n} \to 1$.

10 Show that the root test is more powerful than the ratio test. That is, for any series for which the ratio test implies convergence, the root test must also imply convergence. Moreover, there exist series for which the root test implies convergence, but about which the ratio test gives no information.

11 Suppose $\sum_{n=1}^{\infty} a_n$ is convergent, but not absolutely convergent. Show that for any number c, there is a rearrangement of $\sum_{n=1}^{\infty} a_n$ that converges to c.

12 Suppose $x > 1$. Consider the sequence $x_1, x_2, x_3, \ldots$ defined by setting $x_1 = x$, $x_2 = x^{x_1}$, $x_3 = x^{x_2}$, and in general, $x_{n+1} = x^{x_n}$. It is possible for this sequence to converge? (Use the familiar definition of a^b, $a, b > 0$.)

13 For $n \geq 1$, define

$$s_n = \frac{1}{n}\left[\frac{1}{1 + (1/n)^2} + \cdots + \frac{1}{1 + (n/n)^2}\right].$$

Does the sequence $s_1, s_2, s_3, \ldots$ converge, and if so, to what? (*Hint*: consider the integral $\int_0^1 1/(1 + x^2)\, dx$.)

14 Show that the series $\sum_{n=1}^{\infty} 1/p_n$, where p_n is the nth prime, diverges. (*Hint*: for each k, $1/(1 - 1/p_k) = 1 + 1/p_k + (1/p_k)^2 + \cdots$. Conclude from this that $\{[1 - (1/p_1)]\cdots[1 - (1/p_n)]\}^{-1} > \sum_{k=1}^{n} 1/k$. Since the sum on the right tends to infinity as $n \to \infty$, the product must tend to infinity as $n \to \infty$. The product can be written in the form $[1 + 1/(p_1 - 1)]\cdots[1 + 1/(p_n - 1)]$. Now use the facts that $\sum_{n=1}^{\infty} 1/p_n$ converges if and only if $\sum_{n=1}^{\infty} 1/(p_n - 1)$ converges, and that for $x \geq 0$, $1 + x \leq e^x$.)

15 Show that the Cauchy product of the convergent series $\sum_{n=1}^{\infty} (-1)^{n+1}n^{-1/2}$ with itself diverges.

CHAPTER TWO

Functions

1 Introductory remarks

Suppose S is a subset of the real numbers. By a *function* on S, we mean a rule that assigns some real number to each number in S. Ordinarily we shall use notations such as $f(x)$, $g(x)$, etc., to represent either a function or the value of the function at the point x. This ambiguity should cause no difficulty in practice.

Throughout this book, the set S, called the domain of definition of the function, will usually be an interval. By an interval, we mean a set of numbers satisfying an inequality of the form $a < x < b$, $a \leq x < b$, $a < x \leq b$, or $a \leq x \leq b$, for some numbers a and b with $a \leq b$. (Note that if $a = b$, only the fourth inequality can be satisfied, and the interval reduces to a single point.) The usual notations for the intervals corresponding to the four inequalities are, respectively, (a, b), $[a, b)$, $(a, b]$, and $[a, b]$. An interval of the form (a, b) is called *open*, an interval of the form $[a, b)$ or $(a, b]$ is called *half-open* (or half-closed), and an interval of the form $[a, b]$ is called *closed*. By convention, the symbol $(-\infty, \infty)$ denotes the set of all real numbers, and $[a, \infty)$, (a, ∞), $(-\infty, a)$, and $(-\infty, a]$ denote, respectively, the sets defined by the inequalities $x \geq a$, $x > a$, $x < a$, and $x \leq a$. The notation $x \in I$ is often used to express the fact that the number x is contained in the interval I.

2 Continuous functions

Suppose I is an interval that is either open, closed, or half-open. Suppose $f(x)$ is a function defined on I and x_0 is a point in I. We say that the function $f(x)$ is *continuous* at the point x_0 if the quantity $|f(x) - f(x_0)|$ becomes small as $x \in I$ approaches x_0. More precisely, $f(x)$ is said to be continuous

at $x_0 \in I$ if for any $\epsilon > 0$, there exists a number $\delta > 0$, such that $x \in I$ and $|x - x_0| \leq \delta$ together imply that $|f(x) - f(x_0)| \leq \epsilon$. If $f(x)$ is not continuous at x_0, it is said to be *discontinuous* at x_0.

If $f(x)$ is continuous at every point of an interval, $f(x)$ is said to be continuous on the interval.

Example

It is evident that any function whose value is a constant on an interval I must be continuous on I, since the difference $f(x) - f(x_0)$ must always be 0 for $x_0, x \in I$.

As we shall presently see, any function that is differentiable on an interval I must be continuous on I.

Example

Figure 4 represents the graph of a function $f(x)$ that is *not* continuous on $[0, 1]$, since it is discontinuous at $x = \frac{1}{2}$. Here $f(x) = \frac{1}{2}$ for $x \in [0, \frac{1}{2})$, and $f(x) = 1$ for $x \in [\frac{1}{2}, 1]$.

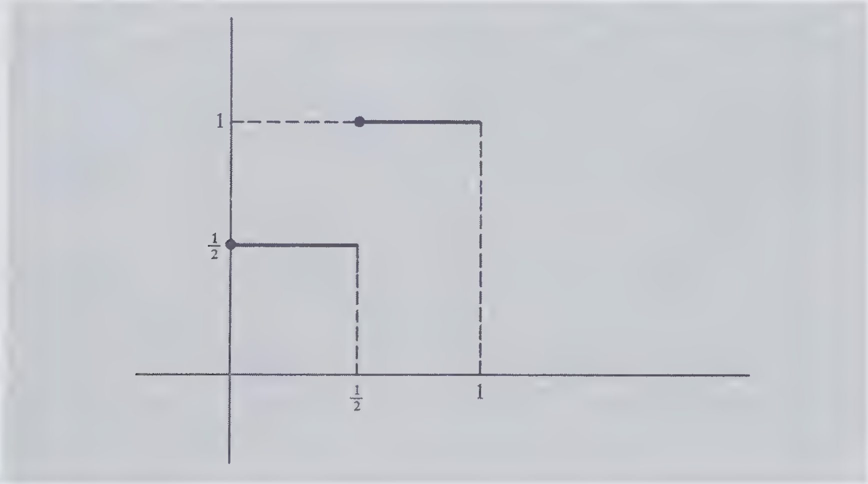

FIGURE 4

We state the following theorem without proof, since the proof is virtually identical to the proofs of Theorems 1.5, 1.6, and 1.7.

Theorem 2.1 *Suppose $f(x)$ and $g(x)$ are continuous on an interval I. Then $f(x) + g(x)$ and $f(x)g(x)$ are both continuous on I. Moreover, if $g(x) \neq 0$ on I, then $f(x)/g(x)$ is continuous on I.*

The following theorem is an immediate consequence of the definition of continuity.

Theorem 2.2 *Suppose I is an interval, and suppose $s_1, s_2, \ldots$ is a convergent sequence of points in I, whose limit s is also a point in I. Then for any function $f(x)$ on I, which is continuous at s, the sequence $f(s_1), f(s_2), \ldots$ is also convergent and converges to $f(s)$.*

There are several important theorems about continuous functions on *closed* intervals. In order to establish them, we require the following preliminary result.

Theorem 2.3 *Suppose $[a, b]$ is a closed interval and $s_1, s_2, \ldots$ is a sequence of numbers in $[a, b]$. Then the sequence $s_1, s_2, \ldots$ has a subsequence $s_1', s_2', \ldots$ that converges to a number $s \in [a, b]$.*

PROOF Denote by S the set of all numbers in $[a, b]$ that stand in the relation $\leq$ to an infinite number of the s_j's. The set S clearly contains at least one element, namely a. Moreover, S is bounded above, since every point in S is $\leq b$. Denote by s the supremum of the set S. Then it is clear that $s \in [a, b]$, and since s is the supremum of the set S, it is easy to see that, for any $\epsilon > 0$, there must be an infinite number of s_j's satisfying $|s - s_j| \leq \epsilon$. From this it follows that there exists a subsequence $s_1', s_2', \ldots$ of $s_1, s_2, \ldots$ having the property that

$$|s - s_1'| \leq 1, |s - s_2'| \leq \tfrac{1}{2}, \ldots, |s - s_n'| \leq \frac{1}{n}, \cdots$$

and hence the sequence $s_1', s_2', \ldots$ converges to s.

Remark The last theorem is not necessarily true for intervals that are not closed. For example, the sequence $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$ does not have a limit in $(0, 1)$, although it converges, of course, to 0.

Theorem 2.4 *Suppose $f(x)$ is continuous on the closed interval $[a, b]$. Then $f(x)$ is bounded on $[a, b]$. That is, there exists a number $M > 0$, such that $|f(x)| \leq M$ for $x \in [a, b]$.*

PROOF Suppose the contrary. Then for each positive integer n there exists a point $s_n \in [a, b]$, such that $|f(s_n)| \geq n$. By Theorem 2.3, the sequence $\{s_n\}$ contains a subsequence $\{s_n'\}$ that converges to a number $s \in [a, b]$. By

Theorem 2.2, $f(s'_n) \to f(s)$ as $n \to \infty$. This is impossible, however, since a convergent sequence is bounded. Hence, we conclude that our assumption is untenable, and consequently, $f(x)$ must be bounded on $[a, b]$.

Remark Theorem 2.4 is false for nonclosed intervals, as the example of the function $f(x) = 1/x$ on $(0, 1)$ shows.

It follows from the last theorem that there is some real number that is the supremum of the set of all numbers of the form $f(x)$, where x ranges over $[a, b]$. Similarly, there is a real number that is the infimum of the set of numbers of the form $f(x)$, where x ranges over $[a, b]$. These quantities are often denoted by $\sup_{x \in [a,b]} f(x)$, and $\inf_{x \in [a,b]} f(x)$, respectively.

Theorem 2.5 *Suppose $f(x)$ is continuous on the closed interval $[a, b]$. Then there exist numbers $x_1, x_2 \in [a, b]$ such that $f(x_1) = \sup_{x \in [a,b]} f(x)$ and $f(x_2) = \inf_{x \in [a,b]} f(x)$.*

PROOF There clearly exists a sequence of numbers $s_1, s_2, \ldots$ in $[a, b]$ having the property that $f(s_n) \to \sup_{x \in [a,b]} f(x)$. By Theorem 2.3, $s_1, s_2, \ldots$ has a subsequence $s'_1, s'_2, \ldots$ that converges to a number $x_1 \in [a, b]$, and it is immediate that $f(x_1) = \sup_{x \in [a,b]} f(x)$. The proof for the remaining case is essentially the same.

Remarks 1 Theorem 2.5 is false for nonclosed intervals, as the example of the function $f(x) = x$ on $(0, 1)$ shows. In this case, $\sup_{x \in (0,1)} f(x) = 1$ and $\inf_{x \in (0,1)} f(x) = 0$, but there is no point in $(0, 1)$ at which $f(x)$ takes on either of these values.

2 According to Theorem 2.5, a continuous function on a closed interval actually takes on its supremum and infimum over the interval. For this reason, these quantities are often referred to as the maximum and minimum, respectively, of the function over the interval.

Theorem 2.6 (Intermediate Value Theorem) *Suppose $f(x)$ is continuous on the closed interval $[a, b]$. Then for any y_0 in the closed interval with endpoints $f(a)$ and $f(b)$, there exists an $x_0 \in [a, b]$, such that $f(x_0) = y_0$. In other words, $f(x)$ takes on all the values between $f(a)$ and $f(b)$.*

PROOF Suppose, for definiteness, that $f(a) \leq f(b)$. (The remaining case can be reduced to this by replacing $f(x)$ by $-f(x)$.) Suppose $y_0 \in [f(a), f(b)]$. Now the midpoint of $[a, b]$ divides $[a, b]$ into two closed subintervals, which overlap at the midpoint. At least one of these intervals must have the property that y_0 is $\leq$ to the value of $f(x)$ at the right endpoint, and $\geq$ to the value of $f(x)$ at the left endpoint. Select such a subinterval, and call it I_1. Now I_1 can be similarly split into two subintervals, at least one of which must have the above-mentioned property. Continuing in this fashion, we obtain a sequence $I_1, I_2, \ldots$ of closed subintervals of $[a, b]$, each contained within the preceding, with lengths tending to zero, and such that for each n, y_0 lies between the values of $f(x)$ at the endpoints of I_n.

Now the sequence $s_1, s_2, \ldots$ of left endpoints of $I_1, I_2, \ldots$ is increasing and bounded above by b and so must converge to a number $x_0 \in [a, b]$. Similarly, the sequence $s_1^*, s_2^*, \ldots$ of right endpoints of $I_1, I_2, \ldots$ is decreasing and bounded below by x_0. Since the length of I_n tends to zero as $n \to \infty$, it is clear that $s_n^* \to x_0$. By continuity, $f(s_n) \to f(x_0)$ and $f(s_n^*) \to f(x_0)$. Since $f(s_n) \leq y_0$ for each n, while $f(s_n^*) \geq y_0$ for each n, we conclude that $y_0 \leq f(x_0) \leq y_0$, which implies that $f(x_0) = y_0$, and this proves the theorem.

Definition 2.1 Suppose a function $f(x)$ is bounded on an interval I. The supremum of the set of all numbers of the form $|f(x) - f(y)|$, with $x, y \in I$, is called the *oscillation* of $f(x)$ over I. It is evident that if I' is a subinterval of I, then the oscillation of $f(x)$ over I' is $\leq$ to the oscillation of $f(x)$ over I.

The following strong form of continuity is often useful.

Definition 2.2 A function $f(x)$ is said to be *uniformly continuous* on an interval I, if for every $\epsilon > 0$, there exists a $\delta > 0$, such that $x, y \in I$ and $|x - y| \leq \delta$ together imply that $|f(x) - f(y)| \leq \epsilon$. Equivalently, $f(x)$ is uniformly continuous over I if for any $\epsilon > 0$, there exists a $\delta > 0$, such that the oscillation of $f(x)$ is $\leq \epsilon$, over any subinterval of I of length $\leq \delta$.

Remark It is clear that uniform continuity over an interval implies continuity on the interval, for if we wish to verify the continuity of a function $f(x)$ at a point $x_0 \in I$, given that $f(x)$ is uniformly continuous on I, we can simply take $y = x_0$ in the definition of uniform continuity, and the result coincides with the definition of the continuity of $f(x)$ at x_0.

On the other hand, continuity of a function $f(x)$ on an interval does not necessarily imply uniform continuity of $f(x)$ on the interval. An example is

furnished by the function $f(x) = \sin(1/x)$ on $(0, 1)$, which is continuous, but not uniformly continuous (see Figure 5).

The next theorem shows that examples like the one described in the preceding remark, of continuous functions that are not uniformly continuous, cannot occur on closed intervals.

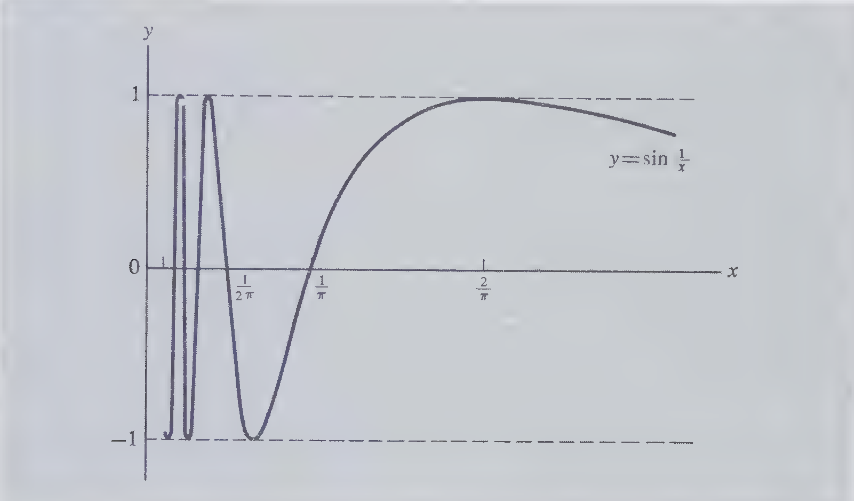

FIGURE 5

Theorem 2.7 *Suppose $f(x)$ is continuous on the closed interval $[a, b]$. Then $f(x)$ is uniformly continuous on $[a, b]$. Thus, for closed intervals, the concepts of continuity and uniform continuity coincide.*

PROOF Suppose $f(x)$ is continuous on $[a, b]$. If $f(x)$ is not uniformly continuous on $[a, b]$, there exists some $\epsilon > 0$ and a sequence $I_1, I_2, \ldots$ of subintervals of $[a, b]$, whose lengths tend to zero, such that the oscillation of $f(x)$ over I_n is $\geq \epsilon$ for each n. Consider the sequence $s_1, s_2, \ldots$, where s_n is the midpoint of I_n. By Theorem 2.3, $s_1, s_2, \ldots$ has a subsequence $s'_1, s'_2, \ldots$, which converges to a point x_0 in $[a, b]$. Now $f(x)$ is continuous at $x = x_0$, so there exists some number $\delta > 0$, such that $x \in I$ and $|x - x_0| \leq \delta$ together imply that $|f(x) - f(x_0)| \leq \frac{1}{4}\epsilon$.

Denote by S, the subinterval of I that consists of all $x \in I$ satisfying $|x - x_0| \leq \delta$. Now if $x \in S$ and $y \in S$, then $|f(x) - f(x_0)| \leq \frac{1}{4}\epsilon$ and $|f(x_0) - f(y)| \leq \frac{1}{4}\epsilon$. By the triangle inequality, this implies that $|f(x) - f(y)| \leq \frac{1}{2}\epsilon$. From this we conclude that the oscillation of $f(x)$ over S is $\leq \frac{1}{2}\epsilon$. Now

for sufficiently large n, I_n must be contained in S. This leads to a contradiction, since the oscillation of $f(x)$ over I_n is $\geq \epsilon$. We conclude that $f(x)$ must be uniformly continuous on $[a, b]$.

3 Sequences and series of functions

Definition 2.3 We say that a *sequence* $f_0(x), f_1(x), f_2(x), \ldots$ of functions defined on an interval I is *convergent* on I, if for each $x_0 \in I$ the sequence of numbers $f_0(x_0), f_1(x_0), f_2(x_0), \ldots$ is convergent.

We say that a *series* $\sum_{n=0}^{\infty} f_n(x)$ of functions defined on I is *convergent* on I, if for each $x_0 \in I$, the series $\sum_{n=0}^{\infty} f_n(x_0)$ is convergent.

The statement that a sequence $\{f_n(x)\}$ converges to a function $f(x)$ is often written $f_n(x) \to f(x)$.

Example

Suppose $f_0(x) = a_0$, $f_1(x) = a_1x$, $f_2(x) = a_2x^2, \ldots$ for some sequence of numbers $a_0, a_1, a_2, \ldots$. The series $\sum_{n=0}^{\infty} a_nx^n$ is called a *power series* (the theory of such series, which is an important branch of mathematics, will be extensively considered in the next chapter). For the time being, we merely note that if $a_0 = a_1 = a_2 = \cdots = 1$, the series becomes the geometric series, which, as we know, converges for $x \in (-1, 1)$.

There is a feature, absent in the ordinary theory of convergent sequences, which complicates the theory of convergent sequences of functions. This is, roughly speaking, that the *rate* at which a sequence $\{f_n(x)\}$ converges on an interval I may be very much dependent upon the value of x. For example, if $f_1(x), f_2(x), \ldots$ is the sequence of functions on $[0, 1]$ for which the graph of the nth function is illustrated in Figure 6, then it is clear that $f_n(x) \to 0$ on $[0, 1]$. The feature of this particular sequence that can be inconvenient for certain purposes, is that for each n there is always a point, in this case $x = 1/2n$, at which the value of the function is n. This might be described by saying that although $f_n(x) \to 0$ on $[0, 1]$, $\{f_n(x)\}$ does not tend to zero *uniformly* on $[0, 1]$.

Definition 2.4 Suppose $\{f_n(x)\}$ is a convergent sequence of functions on an interval I. Denote by $f(x)$ the function to which the sequence converges. Then $\{f_n(x)\}$ is said to *converge uniformly* to $f(x)$ if the function $f(x) - f_n(x)$ is bounded on I for each n, and if $\sup_{x \in I} |f(x) - f_n(x)|$ tends to zero as $n \to \infty$.

A convergent series $\sum_{n=0}^{\infty} f_n(x)$ of functions on I with limit $f(x)$ is said to *converge uniformly* to $f(x)$, if the associated sequence $f_0(x), f_0(x) + f_1(x), \ldots$ of partial sums converges uniformly to $f(x)$.

Example

Suppose $\{f_n(x)\}_{n=0}^{\infty}$ is a bounded sequence of functions on an interval I. That is, suppose there exists $M > 0$, such that for each n, $|f_n(x)| \leq M$, for all $x \in I$. Suppose, also, that $\sum_{n=0}^{\infty} a_n$ is an absolutely convergent series. Then it follows from the majorization test that the series $\sum_{n=0}^{\infty} a_n f_n(x)$ converges on I. Denote by $F(x)$ the function to which the series converges. Then it is easy to see that $\sum_{n=0}^{\infty} a_n f_n(x)$ converges to $F(x)$ uniformly on I, since $|F(x) - \sum_{n=0}^{N} a_n f_n(x)| = |\sum_{n=N+1}^{\infty} a_n f_n(x)| \leq M \sum_{n=N+1}^{\infty} |a_n|$, and the last quantity tends to zero as $N \to \infty$.

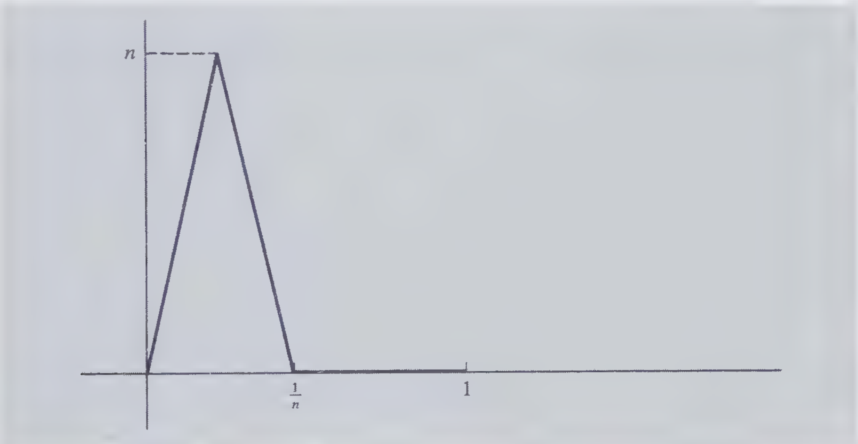

FIGURE 6

An important feature of uniform convergence is described by the following theorem.

Theorem 2.8 *Suppose $\{f_n(x)\}$ is a sequence of continuous functions that converges uniformly to a function $f(x)$ on an interval I. Then $f(x)$ must be continuous on I; in other words, a uniform limit of continuous functions is continuous.*

PROOF Suppose $x_0 \in I$. We wish to show that for any $\epsilon > 0$, there exists a $\delta > 0$, such that $x \in I$ and $|x - x_0| \leq \delta$ together imply that $|f(x) - f(x_0)| \leq \epsilon$. Suppose $\epsilon > 0$ is given. Now it follows from the uniform convergence of $\{f_n(x)\}$ to $f(x)$ that there exists an integer n_0, such that $|f(x) - f_{n_0}(x)| \leq \frac{1}{3}\epsilon$,

for all $x \in I$. Moreover, by the continuity of $f_{n_0}(x)$, there exists a number $\delta > 0$, such that $x \in I$ and $|x - x_0| \leq \delta$ together imply that $|f_{n_0}(x) - f_{n_0}(x_0)| \leq \frac{1}{3}\epsilon$. Now

$$f(x) - f(x_0) = (f(x) - f_{n_0}(x)) + (f_{n_0}(x) - f_{n_0}(x_0)) + (f_{n_0}(x_0) - f(x_0)),$$

so it follows immediately from the triangle inequality that if $x \in I$ and $|x - x_0| \leq \delta$, then $|f(x) - f(x_0)| \leq \epsilon$, which proves the theorem.

Remark The analog of the last theorem is false for nonuniformly convergent sequences of continuous functions. Figure 7 illustrates a situation in which a sequence of continuous functions converges, although not, of course, uniformly, to the discontinuous function on [0, 1] whose value is 0 for $x \in [0, \frac{1}{2})$ and 1 for $x \in [\frac{1}{2}, 1]$.

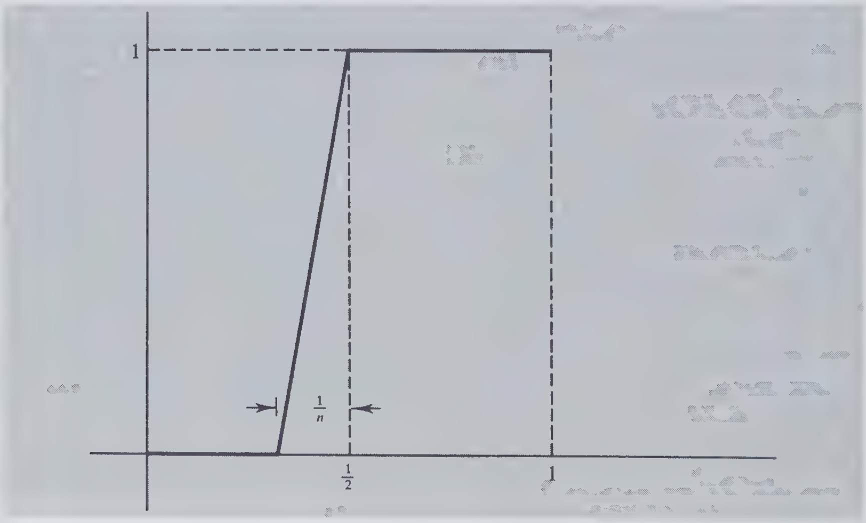

FIGURE 7

4 Other classes of functions

In this section we review very briefly some topics with which we assume the reader is, to some degree, already familiar.

Differentiable functions

Recall that a function $f(x)$, defined on an interval I, is said to be *differentiable* on I if for every $x \in I$, the quotient $(f(y) - f(x))/(y - x)$ tends to a definite

number as y approaches x through values in I that are distinct from x. (If the interval I is closed or half-closed and if x is an endpoint, then y must, of course, approach x from one side.) Equivalently, $f(x)$ is differentiable on I, if for every $x \in I$ there exists a number c, such that if we define a function $g(y)$ on I by setting

$$g(y) = \frac{f(y) - f(x)}{y - x} \qquad \text{if } y \neq x$$

$$g(y) = c \qquad \text{if } y = x,$$

then the function $g(y)$ is continuous at $y = x$.

The number c, which depends on x, is written $f'(x)$, or $f^{(1)}(x)$ and the function $f'(x)$, or $f^{(1)}(x)$, is called the *derivative* of $f(x)$. If the function $f'(x)$ is itself differentiable on I, its derivative is denoted by $f^{(2)}(x)$. More generally, if $f(x)$ has n successive derivatives on I, the nth derivative is denoted by $f^{(n)}(x)$. The function $f(x)$ itself is occasionally denoted by $f^{(0)}(x)$.

It is important to note that if a function is differentiable on an interval I, then it must be continuous on I. For if x is a point in I, we know that the function $g(y)$, which we have defined above, is continuous at $y = x$, and since $f(y) = f(x) + g(y)(y - x)$, this implies that $f(y)$ must be continuous at x.

Definition 2.5 Suppose n is an integer ≥ 1. A function $f(x)$, defined on an interval I, is of *class* C^n on I, if $f(x)$ has at least n successive derivatives on I and if, moreover, the function $f^{(n)}(x)$ is continuous on I. If $f(x)$ is at least continuous on I, it is of *class* C^0 on I, and if $f(x)$ has arbitrarily many derivatives, that is, if $f^{(n)}(x)$ exists for any integer $n \geq 0$, then $f(x)$ is of *class* C^∞ on I. Note that these classes are not mutually exclusive, since clearly any function of class C^n must be of class C^m, for $m \leq n$.

Riemann integrable functions

Suppose $[a, b]$ is a closed interval, and suppose $I_1, I_2, \ldots, I_n$ is a finite collection of nonoverlapping subintervals of $[a, b]$, having the property that $[a, b]$ is the set-theoretic union of $I_1, I_2, \ldots, I_n$. (The I_j's may be open, half-open, or closed.) Then any function $f(x)$ on $[a, b]$, whose value on each I_j is a constant c_j, is called a *step function*. If r_j is the length of I_j, the quantity $r_1c_1 + \cdots + r_nc_n = \sum_{j=1}^{n} r_jc_j$ is called the *integral* of $f(x)$ over I and is designated by the symbol $\int_a^b f(x)\,dx$. (The "dummy variable" in the integral is of no importance, and we could as well write, for example, $\int_a^b f(t)\,dt$.)

Definition 2.6 A function $f(x)$ on an interval $[a, b]$ is said to be *Riemann integrable* over $[a, b]$, if it can be squeezed between two step functions whose integrals are arbitrarily close together. More precisely, $f(x)$ is Riemann

integrable over $[a, b]$, if, for any $\epsilon > 0$, there exist step functions $g_1(x)$ and $g_2(x)$ on $[a, b]$, such that

1. $g_1(x) \leq f(x) \leq g_2(x), \qquad$ for $x \in [a, b]$;
2. $\int_a^b (g_2(x) - g_1(x))\, dx \leq \epsilon$.

It is clear from (1) that a Riemann integrable function must be bounded on $[a, b]$ and that the set of integrals of the step functions smaller than or equal to $f(x)$ is bounded above, and hence has a supremum. This supremum is called the *integral* of $f(x)$. Note that the integral of $f(x)$ could also be defined as the infimum of the set of integrals of the step functions larger than or equal to $f(x)$, since it follows from (2) that the latter quantity exists and must equal the value given by the previous definition of the integral. The integral of a Riemann integrable function $f(x)$ is denoted by $\int_a^b f(x)\, dx$, and the symbol $\int_b^a f(x)\, dx$, by definition, denotes the quantity $-\int_a^b f(x)\, dx$.

We list a few well-known properties of the integral.

1. $\displaystyle\int_a^b cf(x)\, dx = c\int_a^b f(x)\, dx.$

2. $\displaystyle\int_a^b (f(x) + g(x))\, dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx.$

3. $\displaystyle\int_a^b f(x)\, dx + \int_b^c f(x)\, dx = \int_a^c f(x)\, dx.$

4. $\displaystyle\int_b^a f(x)\, dx = -\int_a^b f(x)\, dx.$

5. If $f(x) \geq g(x)$ on $[a, b]$, then $\displaystyle\int_a^b f(x)\, dx \geq \int_a^b g(x)\, dx.$

6. If $|f(x)| \leq M$ on $[a, b]$, then $\displaystyle\left|\int_a^b f(x)\, dx\right| \leq M(b - a).$

7. If $f(x)$ is continuous on $[a, b]$, then $f(x)$ is Riemann integrable over $[a, b]$.

The following very important property, called the *fundamental theorem of calculus*, establishes the link between the derivative and the integral.

8. Suppose $f(x)$ is continuous on $[a, b]$. Define a function $F(x)$ on $[a, b]$, by setting $F(x) = \int_a^x f(t)\,dt$. Then $F(x)$ is of class C^1 on $[a, b]$, and $F'(x) = f(x)$.

We shall prove (7).

By Theorem 2.7, $f(x)$ is uniformly continuous on $[a, b]$. Thus, for any $\epsilon > 0$, there exists a $\delta > 0$, such that on any subinterval of $[a, b]$ of length $\leq \delta$, the oscillation of $f(x)$ is $\leq \epsilon/(b - a)$. Suppose $I_1, I_2, \ldots, I_n$ is a partition of $[a, b]$ into nonoverlapping subintervals, each of length $\leq \delta$. Define $C_j = \sup_{x \in I_j} f(x)$ and $c_j = \inf_{x \in I_j} f(x)$. Then if we define step functions $g_1(x)$ and $g_2(x)$, so that $g_1(x)$ has the value c_j on I_j, and $g_2(x)$ has the value C_j on I_j, it is evident that $g_1(x)$ and $g_2(x)$ satisfy conditions (1) and (2) in the definition of the integral, which proves (7).

The following result will be useful in the next chapter.

Theorem 2.9 *Suppose $\{f_n(x)\}_{n=0}^{\infty}$ is a sequence of functions, each of which is Riemann integrable over an interval $[a, b]$. Suppose, moreover, that the sequence $\{f_n(x)\}_{n=0}^{\infty}$ converges uniformly on $[a, b]$ to a function $f(x)$, which is again Riemann integrable over $[a, b]$. Then the associated sequence of integrals*

$$\int_a^b f_0(x)\,dx, \int_a^b f_1(x)\,dx, \int_a^b f_2(x)\,dx, \ldots \textit{ converges to } \int_a^b f(x)\,dx.$$

PROOF It suffices to show that $\int_a^b (f(x) - f_n(x))\,dx \to 0$. Now for any $\epsilon > 0$, there exists an integer n_0, such that for $n \geq n_0$, $|f(x) - f_n(x)| \leq$ $\leq \epsilon/(b - a)$, for $x \in [a, b]$. Thus, for $n \geq n_0$,

$$\left| \int_a^b (f(x) - f_n(x))\,dx \right| \leq \frac{\epsilon}{(b - a)}(b - a) = \epsilon,$$

and this establishes the theorem.

Corollary *If $\sum_{n=0}^{\infty} f_n(x)$ is a series of Riemann integrable functions on $[a, b]$ which converges uniformly to a Riemann integrable function $f(x)$ on $[a, b]$, then the series $\sum_{n=0}^{\infty} (\int_a^b f_n(x)\,dx)$ converges to $\int_a^b f(x)\,dx$.*

Exercises

1 Define a function $f(x)$ on $(0, 1)$ by setting $f(x) = 1/q$, if x is a rational number whose expression in lowest terms is p/q (p, q positive) and $f(x) = 0$ if x is irrational. At which points, if any, of $(0, 1)$ is $f(x)$ continuous?

2 Show that if $P(x) = a_n x^n + \cdots + a_1 x + a_0$ is a polynomial for which the a_j's are real numbers, with $a_n \neq 0$ and n odd, then there must exist a real number x_0, such that $P(x_0) = 0$. Is this necessarily true if n is even?

3 Suppose $\{f_n(x)\}$ is a sequence of continuous functions on $[0, 1]$ that converges, though not necessarily uniformly, to a continuous function $f(x)$ on $[0, 1]$. Is it necessarily true that

$$\int_0^1 f_n(x)\,dx \to \int_0^1 f(x)\,dx?$$

4 Does there exist a bounded function on $[0, 1]$ that is not Riemann integrable over $[0, 1]$?

5 A function $f(x)$, defined on a symmetric interval about the origin, is said to be *even* if $f(-x) = f(x)$, while $f(x)$ is said to be *odd* if $f(-x) = -f(x)$. For example, the function whose value at x is x^2 is even, while the function whose value at x is x^3 is odd. Show that any function defined on a symmetric interval about the origin can be expressed as the sum of an even function and an odd function. Can there be more than one way of doing this? That is, if $f(x) = g(x) + h(x) = g^*(x) + h^*(x)$, where $g(x)$ and $g^*(x)$ are even and $h(x)$ and $h^*(x)$ are odd, must $h(x) = h^*(x)$ and $g(x) = g^*(x)$?

6 Show that the derivative of a differentiable odd function must be even.

7 Suppose $f(x)$ is continuous on an interval I, and suppose $g(x)$ is continuous on an interval containing all numbers of the form $f(x_0)$, with x_0 in I. Show that $g(f(x))$ is continuous on I.

8 Suppose $f(x)$ is a continuous function on $[0, 1]$ that takes $[0, 1]$ into itself. That is, suppose $0 \leq f(x) \leq 1$ for x in $[0, 1]$. Show that there must exist some number x_0 in $[0, 1]$, such that $f(x_0) = x_0$. If we replace $[0, 1]$ by $(0, 1)$, is the corresponding statement true?

9 Suppose $f(x)$ is a continuous function on $(-\infty, \infty)$ that satisfies the equation $f(a + b) = f(a) + f(b)$, for any two numbers a and b. Show that $f(x)$ must be of the form cx, for some constant c. (*Hint*: define $c = f(1)$. Show that the function $g(x) = f(x) - cx$ satisfies $g(p/q) = 0$, for any integers p and q with $q \neq 0$.)

10 Suppose $f(x)$ is continuous on $[0, \infty)$, and suppose that for any $x_0 > 0$, $f(nx_0) \to 0$ as $n \to \infty$. Is it necessarily true that $f(x) \to 0$ as $x \to \infty$?

CHAPTER THREE

Power Series

1 Expansions about the origin

One of the most useful ideas in mathematics is that of expressing, when possible, a given function in terms of simpler functions whose properties are well understood. There are many examples of this; one of the most natural and important is furnished by the so-called power series, which we have encountered before. From another point of view, power series can, of course, also be investigated for their own sake, without reference to any desire to represent a given function in terms of them, and we shall begin our discussion with this latter aim in mind.

Suppose we are given a power series $\sum_{n=0}^{\infty} a_n x^n$. There are several questions we can then ask: one of the most basic is: for what values of x does the series converge? It is worth noting at this point that the series always converges at the origin, that is, for $x = 0$, it converges to the value a_0. What more can we say in general? The following lemma will be useful in the investigation of this question.

Lemma 3.1 *Suppose $\sum_{n=0}^{\infty} a_n x^n$ is a power series, and suppose that for some x_0 it is known that the terms of the series are bounded. In other words, suppose that there exists an $M > 0$ such that $|a_n x_0^n| \leq M$ $(n = 0, 1, 2, \ldots)$. Then the series must be absolutely convergent for all x such that $|x| < |x_0|$.*

PROOF Suppose x satisfies $|x| < |x_0|$. Now $a_n x^n = a_n x_0^n (x/x_0)^n$, so $|a_n x^n| \leq M|x/x_0|^n$. But $|x/x_0| < 1$, so the series $\sum_{n=0}^{\infty} a_n x^n$ is majorized by the absolutely convergent geometric series $\sum_{n=0}^{\infty} Mr^n$, where $r = |x/x_0|$, and hence is itself absolutely convergent.

Definition 3.1 The *radius of convergence* of a power series $\sum_{n=0}^{\infty} a_n x^n$ is the supremum of the set of all x's for which $\sum_{n=0}^{\infty} |a_n x^n|$ converges. Since

any power series converges for $x = 0$, this set is nonempty. There are thus three possibilities. If we denote the radius of convergence of a given series by r, then either $r = 0$, r is some positive real number, or $r = \infty$. (Recall that the last statement is simply an abbreviated way of expressing the fact that the set of x's for which the series converges absolutely is not bounded above.)

Theorem 3.1 *Suppose $\sum_{n=0}^{\infty} a_n x^n$ is a power series with radius of convergence r. Then*

1. *If $r = 0$, $\displaystyle\sum_{n=0}^{\infty} a_n x^n$ converges only for $x = 0$.*
2. *If r is a positive real number, $\displaystyle\sum_{n=0}^{\infty} a_n x^n$ is absolutely convergent for all x satisfying $|x| < r$ and diverges for all x satisfying $|x| > r$.*
3. *If $r = \infty$, $\displaystyle\sum_{n=0}^{\infty} a_n x^n$ is absolutely convergent for all x.*

Remark In statement (2) we assert that the series converges absolutely in the interval $(-r, r)$ and diverges outside of the interval $[-r, r]$. This leaves the status of the endpoints $x = -r$ and $x = r$ in doubt. The question of convergence, or lack of it, at the endpoints is, in general, delicate and must be settled by individual inspection in each case.

PROOF Statement (3) is obvious. In order to prove statements (1) and (2), it clearly suffices to show that $\sum_{n=0}^{\infty} a_n x^n$ diverges for all x's satisfying $|x| > r$. Suppose it does not diverge. Then there is an x_0 such that $|x_0| > r$ and $\sum_{n=0}^{\infty} a_n x_0^n$ converge. Since $\sum_{n=0}^{\infty} a_n x_0^n$ converges, its terms must tend to zero and, in particular, must be bounded. By Lemma 3.1, this implies that $\sum_{n=0}^{\infty} a_n x^n$ must be absolutely convergent for all x satisfying $|x| < |x_0|$, which contradicts the definition of r.

We shall now describe various ways of determining the radius of convergence of a series.

The Ratio Method Suppose $\sum_{n=0}^{\infty} a_n x^n$ is a power series for which from some point onward none of the a_n's are zero. That is, for some n_0, $a_n \neq 0$ for $n \geq n_0$. Suppose we form the sequence of ratios

$$\left|\frac{a_{n_0+1}}{a_{n_0}}\right|, \quad \left|\frac{a_{n_0+2}}{a_{n_0+1}}\right|, \quad \left|\frac{a_{n_0+3}}{a_{n_0+2}}\right|, \ldots,$$

and suppose this sequence tends to a limit L. Then $r = 1/L$. If $L = 0$, we interpret this to mean that the series converges for all x, which we have agreed to indicate by $r = \infty$. Finally, if the sequence tends to infinity, $r = 0$.

PROOF The result follows almost immediately from the ratio test of Chapter 1. To begin with, suppose that L is finite and not equal to 0. Then if $|x| < 1/L$, the ratio $|a_{n+1}x^{n+1}/a_nx^n| = |a_{n+1}x/a_n|$ is, beyond some point, less than $1 - \epsilon$, for some $\epsilon > 0$, and the series converges. Similarly, if $|x| > 1/L$, the ratio is greater than $1 + \epsilon$, and the series diverges, since its terms do not tend to zero. If $L = 0$, it is clear that for any x, the ratio $|a_{n+1}x/a_n|$ is, beyond some point, less than $1 - \epsilon$, and the series converges for all x. Finally, if $|a_{n+1}/a_n| \to \infty$, it is clear that for any $x \neq 0$, the ratio $|a_{n+1}x/a_n|$ is, beyond some point, greater than $1 + \epsilon$, and the series diverges for any $x \neq 0$.

The Root Method Suppose $\sum_{n=0}^{\infty} a_nx^n$ is a series for which $L = \lim_{n\to\infty} |a_n|^{1/n}$ exists. Then

a. $r = \dfrac{1}{L}$, if $L > 0$;

b. $r = \infty$, if $L = 0$;

c. $r = 0$, if $|a_n|^{1/n} \to \infty$.

PROOF This follows immediately from the root test of Chapter 1, since $|a_nx^n|^{1/n} = |x|\,|a_n|^{1/n}$.

Examples

1. $\sum_{n=0}^{\infty} x^n/n!$. The ratio test shows that $r = \infty$ for this series. (The root test could be applied to give the same result, but the ratio test is more convenient here.)

2. $\sum_{n=0}^{\infty} x^n/n^n$. The root test shows that $r = \infty$ for this series.

One difficulty of applying either of these methods stems from the fact that one or the other of the limits $\lim_{n\to\infty} |a_{n+1}/a_n|$ and $\lim_{n\to\infty} |a_n|^{1/n}$ may not exist. A case in point is the series

$$1 + 2x + x^2 + 2^3x^3 + x^4 + 2^5x^5 + \cdots,$$

in which the coefficient of x^n for n even is 1, while the coefficient of x^n for n odd is 2^n. Thus in this case the ratio method is useless, while the quantities $|a_n|^{1/n}$ that occur in the root method oscillate between 1 and 2 without approaching a definite limit. It is not difficult to show that for this series the

radius of convergence is $\frac{1}{2}$, but neither of the two methods we have described is adequate for the purpose. Of the two, the root method provides greater insight, and we intuitively suspect, and in fact we can prove, that the radius of convergence must be the reciprocal of the larger of the two quantities between which $|a_n|^{1/n}$ oscillates. In a certain sense, this rather vague principle is always true, as the following theorem shows.

Theorem 3.2 (Cauchy-Hadamard) *Given a series* $\sum_{n=0}^{\infty} a_n x^n$, *define* $L = \overline{\lim}\, |a_n|^{1/n}$, *where by* $\overline{\lim}\, |a_n|^{1/n}$, *we mean the supremum of the set of numbers that are less than an infinite number of the* $|a_n|^{1/n}$'s. *Then*

a. $r = \dfrac{1}{L}$, *if* $L > 0$;

b. $r = \infty$, *if* $L = 0$;

c. $r = 0$, *if* $L = \infty$, *that is, if the* $|a_n|^{1/n}$'s *are not bounded.*

PROOF Suppose L is finite and nonzero. Then it follows from the definition of L that for sufficiently small $\epsilon > 0$, the terms of the series $\sum_{n=0}^{\infty} |a_n x^n|$ are, beyond some point, dominated by those of the series $\sum_{n=0}^{\infty} (L + \epsilon)^n |x|^n$, while, on the other hand, an infinite number of terms of the series $\sum_{n=0}^{\infty} |a_n x^n|$ are larger than the corresponding terms of the series $\sum_{n=0}^{\infty} (L - \epsilon)^n |x|^n$. By the root test, $\sum_{n=0}^{\infty} (L + \epsilon)^n |x|^n$ converges for $|x| < 1/(L + \epsilon)$, while the terms of $\sum_{n=0}^{\infty} (L - \epsilon)^n |x|^n$ are not bounded for $|x| > 1/(L - \epsilon)$. This implies that $\sum_{n=0}^{\infty} a_n x^n$ converges for $|x| < 1/(L + \epsilon)$, and diverges for $|x| > 1/(L - \epsilon)$. Since ϵ can be arbitrarily small, this shows that $r = 1/L$. If $L = 0$, the above argument shows that $r = \infty$, while if $L = \infty$ and $x \neq 0$, the terms of the series $\sum_{n=0}^{\infty} |a_n x^n|$ are not even bounded, since their nth roots are not bounded; thus, $r = 0$ in this case.

So far, we have said nothing about the nature of the function that a given power series represents within its interval of convergence. As we shall see, such a function must perforce be of class C^∞ within the interior of the interval of convergence of the power series. We shall also see that there exist C^∞ functions that cannot be expressed by power series, so that the class of functions that have power series expansions in a given interval forms a proper subclass of the class of functions that are C^∞ within that interval.

In general, we call a function *analytic* in an interval $(-r, r)$, if it is defined in $(-r, r)$ and represented there by a power series. More precisely, $f(x)$ is analytic if there exists a power series $\sum_{n=0}^{\infty} a_n x^n$, whose radius of convergence is at least r, and for which $f(x) = \sum_{n=0}^{\infty} a_n x^n$ for x in $(-r, r)$. (NOTE: the term "analytic" is often used to denote a slightly different concept, namely,

that of a so-called real-analytic function in an interval. In the context of this book, our usage will be unambiguous.)

To return to our discussion, suppose $\sum_{n=0}^{\infty} a_n x^n$ is a power series with radius of convergence $r > 0$ and denote by $f(x)$ the function to which it converges in $(-r, r)$. (NOTE: the series $\sum_{n=0}^{\infty} a_n x^n$ may, of course, be convergent at either or both of the endpoints, but in what follows, we shall not be concerned with the behavior of $f(x)$ at the endpoints.) Suppose now x_0 is a number in $(-r, r)$. Then $\sum_{n=0}^{\infty} |a_n x_0^n|$ converges. Moreover, if x_1 is such that $|x_1| \leq |x_0|$, then the terms of the series $\sum_{n=0}^{\infty} |a_n x_1^n|$ are dominated by those of the series $\sum_{n=0}^{\infty} |a_n x_0^n|$. From this we conclude that $\sum_{n=0}^{\infty} a_n x^n$ converges uniformly on $[-|x_0|, |x_0|]$. Since a uniform limit of continuous functions is continuous, we conclude that $f(x)$ is continuous on $[-|x_0|, |x_0|]$. Since x_0 can be any point in $(-r, r)$, we conclude that $f(x)$ is continuous in $(-r, r)$.

Much more can be said about $f(x)$. We see this, by noting first that the series $a_1 + 2a_2 x + 3a_3 x^2 + 4a_4 x^3 + \cdots$, which is obtained from $\sum_{n=0}^{\infty} a_n x^n$ by formal differentiation, has the same radius of convergence as $\sum_{n=0}^{\infty} a_n x^n$. This follows from the Cauchy-Hadamard theorem, and from the fact that $\overline{\lim}\, |na_n|^{1/n} = \overline{\lim}\, n^{1/n}|a_n|^{1/n} = \overline{\lim}\, |a_n|^{1/n}$, since $n^{1/n} \to 1$ as $n \to \infty$.

Thus, by the preceding reasoning, the function $g(x)$, represented by the series $a_1 + 2a_2 x + 3a_3 x^2 + 4a_4 x^3 + \cdots$ is continuous in $(-r, r)$, and moreover, the last series converges uniformly to $g(x)$ on any closed subinterval of $(-r, r)$. Since the sum of the integrals of a uniformly convergent series of continuous functions is the integral of the sum, we conclude that for any x in $(-r, r)$,

$$\int_0^x g(t)\, dt = \int_0^x a_1\, dt + \int_0^x 2a_2 t\, dt + \int_0^x 3a_3 t^2\, dt + \cdots,$$

or

$$\begin{aligned} \int_0^x g(t)\, dt &= a_1 x + a_2 x^2 + a_3 x^3 + \cdots \\ &= f(x) - a_0. \end{aligned}$$

This shows that $f(x)$ is differentiable in $(-r, r)$, and moreover, $f'(x) = g(x)$.

Thus, since $g(x)$ is continuous in $(-r, r)$, we have proved that $f(x)$ is C^1 in $(-r, r)$. If, now, we repeat the above chain of reasoning for $g(x)$, we find that $g(x)$ itself is C^1 in $(-r, r)$, and hence $f(x)$ is C^2 in $(-r, r)$. By induction, we obtain the following important theorem.

Theorem 3.3 *Suppose the series $\sum_{n=0}^{\infty} a_n x^n$ has radius of convergence $r > 0$, and suppose $f(x)$ is the function to which the series converges. Then $f(x)$ is C^∞ in $(-r, r)$. Moreover, the power series obtained from $\sum_{n=0}^{\infty} a_n x^n$ by n differentiations has radius of convergence r, and converges to $f^{(n)}(x)$.*

Example

Theorem 3.3 can sometimes be used to find a simple expression for the function represented by a given power series. Consider the series $x + x^2/2 + x^3/3 + \cdots$. It follows from the root test that the radius of convergence of this series is 1. Moreover, if we designate by $f(x)$ the function to which the series converges in $(-1, 1)$, it is clear from the last theorem that

$$\begin{aligned} f'(x) &= 1 + x + x^2 + x^3 + \cdots \\ &= \frac{1}{1 - x} \end{aligned}$$

Furthermore, since $f(0) = 0$, we must have $f(x) = \int_0^x f'(t)\,dt$ for $x \in (-1, 1)$. That is, for $x \in (-1, 1)$, $f(x) = -\log(1 - x)$.

2 Algebraic operations with power series

Theorem 3.4 *Suppose $\sum_{n=0}^{\infty} a_n x^n$ is a power series, and suppose that $f(x)$ is the function to which it converges on its interval of convergence. Then for any constant c, the power series $\sum_{n=0}^{\infty} ca_n x^n$ converges to $cf(x)$ on the interval of convergence of the original series, and, if $c = 0$, it converges to 0 on the whole line.*

PROOF This follows immediately from Theorem 1.4.

Theorem 3.5 *Suppose $\sum_{n=0}^{\infty} a_n x^n$ and $\sum_{n=0}^{\infty} b_n x^n$ are two power series. Denote by $f(x)$ and $g(x)$ the functions to which the series converge on their respective intervals of convergence. Then the power series $\sum_{n=0}^{\infty} (a_n + b_n)x^n$ converges to $f(x) + g(x)$, for all values of x for which the two original series are simultaneously convergent. Similarly, the power series $\sum_{n=0}^{\infty} (a_n - b_n)x^n$ converges to $f(x) - g(x)$, for all values of x for which the two original series are simultaneously convergent.*

PROOF The proof follows immediately from Theorem 1.5 and its Corollary.

Remark It may very well happen that one of the series $\sum_{n=0}^{\infty} (a_n + b_n)x^n$, $\sum_{n=0}^{\infty} (a_n - b_n)x^n$, converges at points where neither of the two original series converge. An extremely simple example of this occurs when $a_n = b_n = 1$, for all $n \geq 0$. Then $\sum_{n=0}^{\infty} (a_n - b_n)x^n$ converges to 0 for all x, while the two original series diverge for $|x| \geq 1$.

Suppose now that $\sum_{n=0}^{\infty} a_n x^n$ and $\sum_{n=0}^{\infty} b_n x^n$ are two power series with respective radii of convergence r_1 and r_2. Denote by $f(x)$ and $g(x)$ the functions to which the respective series converge in $(-r_1, r_1)$ and $(-r_2, r_2)$. Then the product $h(x) = f(x)g(x)$ is certainly defined in the interval $|x| < \min(r_1, r_2)$, and we may ask whether or not $h(x)$ has a power series representation in that interval. If such a series exists at all, it is, at least from an intuitive point of view, fairly clear what the coefficients of the series must be. Namely, if we formally multiply $\sum_{n=0}^{\infty} a_n x^n$ by $\sum_{n=0}^{\infty} b_n x^n$, we obtain an infinite array of terms, which can be conveniently collected in the following way:

$$\begin{array}{ccccc}
a_0 b_0 & a_1 b_0 x & a_2 b_0 x^2 & a_3 b_0 x^3 & \cdots \\
a_0 b_1 x & a_1 b_1 x^2 & a_2 b_1 x^3 & a_3 b_1 x^4 & \cdots \\
a_0 b_2 x^2 & a_1 b_2 x^3 & a_2 b_2 x^4 & a_3 b_2 x^5 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \cdots
\end{array}$$

If we now proceed on a purely formal basis and collect the terms containing x^n, we are led to conjecture that the coefficient of x^n in the power series for $h(x)$, if it has one, must be $a_0 b_n + a_1 b_{n-1} + \cdots + a_{n-1} b_0$. This reasoning can be completely justified, as the following theorem shows.

Theorem 3.6 *Suppose $r_1, r_2, f(x), g(x)$, and $h(x)$ are defined as above. Then $h(x)$ has a power series representation $\sum_{n=0}^{\infty} c_n x^n$, where $c_n = \sum_{k=0}^{n} a_k b_{n-k}$. The radius of convergence of $\sum_{n=0}^{\infty} c_n x^n$ is always $\geq \min(r_1, r_2)$.*

PROOF The proof follows immediately from Theorem 1.20. For suppose x_0 is such that $|x_0| < \min(r_1, r_2)$. Then $\sum_{n=0}^{\infty} a_n x_0^n$ and $\sum_{n=0}^{\infty} b_n x_0^n$ are absolutely convergent. Thus, by Theorem 1.20, the Cauchy product of the two series converges to the desired value.

Remark As in the case of sums, it can happen that the radius of convergence of the series for the product is greater than $\min(r_1, r_2)$. For example, the radius of convergence of the series $1 - x$ is ∞, while that for the series $1 + x + x^2 + x^3 + \cdots = 1/(1 - x)$ is 1, so in this case, $\min(r_1, r_2) = 1$. On the other hand, the only nonvanishing term of the product series is the constant term 1, so the product series has radius of convergence ∞.

So far, we have shown that sums and products of functions representable by power series are, in the sense of the last two theorems, also representable

by power series. The next natural question is whether or not quotients have this property. To answer this, it suffices to settle the question of when $1/f(x)$ is representable by a power series, given that $f(x)$ is so representable. If we attempt to answer this last question, it is at once apparent that we must place some restrictions on $f(x)$. In particular, if we want $1/f(x)$ to be represented in an interval $(-r, r)$, it is clear that $f(x)$ must not vanish at any point of $(-r, r)$. As it turns out, this requirement, by itself, is not sufficient. A clear understanding of why this is so, and an illuminating viewpoint for much of the material of this section, is provided by the theory of functions of a complex variable. We shall say more about this at the conclusion of this chapter. For the time being, we state the following result.

Theorem 3.7 *Suppose $f(x)$ has a power series representation in some open interval about the origin, and suppose $f(0) \neq 0$. Then $1/f(x)$ has a power series representation in some open interval about the origin.*

Remark It is evident from the continuity of $f(x)$ and from the fact that $f(x)$ is not zero at the origin that $1/f(x)$ is defined in some open interval about the origin. Suppose now that the radius of convergence of the power series that represents $f(x)$ is r, and denote by r_1 the largest number $\leq r$, for which $f(x) \neq 0$ in $(-r_1, r_1)$. Then it is evident that $1/f(x)$ is defined in $(-r_1, r_1)$, and the temptation is then to suppose that the power series for $1/f(x)$ has a radius of convergence r_1. This is not generally true, and the theorem, if read carefully, does not claim this much. It only asserts that there is some number r_2, which could quite possibly be smaller than r_1, such that $1/f(x)$ has a power series representation in $(-r_2, r_2)$.

We shall not prove Theorem 3.7 above immediately in this section, but we shall show that it follows naturally from the answer to the following general question: given that $f(x)$ and $g(x)$ can be expanded in power series, when can one infer that $f(g(x))$ has a power series expansion? To answer this question, certain restrictions are obviously necessary. For one thing, it is evident that the range of values taken on by $g(x)$ must be contained in the domain of definition of $f(x)$. The whole question is answered by the following theorem.

Theorem 3.8 *Suppose that $f(x)$ and $g(x)$ are analytic in some open interval about the origin, and suppose $g(0) = 0$. Then the function $f(g(x))$ is analytic in some (possibly different) open interval about the origin.* [Note that the function $f(g(x))$ is clearly defined for x in a small enough interval about the origin, since $g(x)$ is continuous.]

PROOF Suppose $r_1 > 0$ is small enough so that $f(x)$ is analytic in $(-r_1, r_1)$, and suppose $r_2 > 0$ is small enough so that $g(x)$ is analytic in $(-r_2, r_2)$, and $|g(x)| < r_1$ for $x \in (-r_2, r_2)$. Then the expression $f(g(x))$ has a definite meaning for x in $(-r_2, r_2)$. Now it follows from Theorem 3.6 that the powers $g(x), (g(x))^2, (g(x))^3, \ldots$ are all analytic in $(-r_2, r_2)$. Accordingly, let $\sum_{n=1}^{\infty} b_{kn}x^n$ be the power series for $(g(x))^k$. (Since $g(0) = 0$, we can start the summation at $n = 1$ for any k.) Each of the above power series converges in $(-r_2, r_2)$, and so, for a fixed $x \in (-r_2, r_2)$, it represents a definite number. Now suppose $\sum_{n=0}^{\infty} a_n x^n$ is the power series for $f(x)$. Then for $x \in (-r_2, r_2)$, the series $\sum_{n=0}^{\infty} y_n$ converges to $f(g(x))$, where

$$\begin{aligned} y_0 &= a_0 \\ y_1 &= a_1(b_{11}x + b_{12}x^2 + b_{13}x^3 + \cdots) \\ y_2 &= a_2(b_{21}x + b_{22}x^2 + b_{23}x^3 + \cdots) \\ y_3 &= a_3(b_{31}x + b_{32}x^2 + b_{33}x^3 + \cdots) \\ &\vdots \end{aligned} \tag{3.1}$$

Thus, if it were possible to apply Fubini's theorem to (3.1), column summation would yield

$$f(g(x)) = c_0 + c_1x + c_2x^2 + \cdots,$$

where

$$\begin{aligned} c_0 &= a_0 \\ c_1 &= a_1b_{11} + a_2b_{21} + a_3b_{31} + \cdots \\ c_2 &= a_1b_{12} + a_2b_{22} + a_3b_{32} + \cdots \\ &\vdots \end{aligned}$$

and this would show that $f(g(x))$ could be expanded in a power series in $(-r_2, r_2)$.

Unfortunately, the application of Fubini's theorem requires that the array on the right side of (3.1) continue to be row summable if we replace each entry by its absolute value. This is not necessarily the case, at least not for all x in $(-r_2, r_2)$, or, what amounts to the same thing, for all $x \in [0, r_2)$. The reason for the difficulty is that not all the a's and b's in (3.1) are necessarily ≥ 0. This inconvenience is only minor, however, and we can get around it as follows.

Consider the series $\sum_{n=0}^{\infty} |a_n|x^n$, which converges to an analytic function $f^*(x)$ in $(-r_1, r_1)$, and the series $\sum_{n=1}^{\infty} |b_{1n}|x^n$, which converges to an analytic function $g^*(x)$ in $(-r_2, r_2)$. Now $g^*(0) = 0$, so by continuity, there exists an interval $(-r_3, r_3)$ such that $|g^*(x)| < r_1$ for $x \in (-r_3, r_3)$. Moreover, as

before, the powers of $g^*(x)$ can be expanded within $(-r_3, r_3)$ in series of the form $(g^*(x))^k = \sum_{n=1}^{\infty} b_{kn}^* x^n$, with $b_{kn}^* \geq |b_{kn}|$, for all combinations of k and n.

By the reasoning that led to (3.1), we find that for $x \in (-r_3, r_3)$

$$f^*(g^*(x)) = \sum_{n=0}^{\infty} y_n^*,$$

where

$$\begin{aligned} y_0^* &= |a_0| \\ y_1^* &= |a_1|(b_{11}^* x + b_{12}^* x^2 + b_{13} x^3 + \cdots) \\ y_2^* &= |a_2|(b_{21}^* x + b_{22}^* x^2 + b_{23} x^3 + \cdots) \\ y_3^* &= |a_3|(b_{31}^* x + b_{32}^* x^2 + b_{33}^* x^3 + \cdots) \\ &\vdots \end{aligned} \tag{3.2}$$

Now the right side of (3.2) satisfies the conditions for Fubini's theorem, and since $b_{kn}^* \geq |b_{kn}|$, it follows that Fubini's theorem can be applied to the right side of (3.1), provided x is in $(-r_3, r_3)$, and this proves the theorem.

We shall now show how the theorem about reciprocals follows immediately from Theorem 3.8.

PROOF OF THEOREM 3.7 Suppose $h(x)$ is analytic in an interval about the origin, and suppose $h(0) = c \neq 0$. We wish to show that $1/h(x)$ is then analytic in an interval about the origin. Note that there is no loss of generality if we suppose $c = 1$; for in the contrary case, we can simply replace $h(x)$ by $(1/c)h(x)$ and prove the theorem for the latter function, which clearly implies the desired result for the original function.

Now define a function $g(x)$ by setting $g(x) = 1 - h(x)$. Note that $g(x)$ is analytic in an interval about the origin, and $g(0) = 0$. Moreover, $h(x) = 1 - g(x)$. Consider next the function

$$\begin{aligned} f(x) &= \frac{1}{1-x} \\ &= 1 + x + x^2 + x^3 + \cdots \end{aligned}$$

which is analytic in $(-1, 1)$. By Theorem 3.8, $f(g(x))$ is analytic in an interval about the origin. That is, $1/(1 - g(x)) = 1/h(x)$ is analytic in an interval about the origin, which was to be proved.

3 Taylor's theorem

We now consider the following question. Suppose a function $f(x)$ is given, by some method other than by a power series, in a neighborhood of the origin; when can $f(x)$ be expanded in a power series about the origin? In other words, when is $f(x)$ analytic in a neighborhood about the origin? We have already encountered isolated instances of this problem. For example, the function $1/(1 - x)$ is analytic in $(-1, 1)$ and has the power series $1 + x + x^2 + x^3 + \cdots$. By integrating the latter, we found that the function $\log(1 - x)$ is also analytic in $(-1, 1)$ and has the power series $-x - x^2/2 - x^3/3 - x^4/4 - \cdots$.

What, however, can be said about the problem in general, without invoking special devices, specifically adapted to the particular problem at hand. In order to gain insight into the general question, let us suppose for the moment that we are given a function $f(x)$, defined in a neighborhood $(-r, r)$ of the origin and known in advance to be analytic in $(-r, r)$.

Suppose $\sum_{n=0}^{\infty} a_n x^n$ is the power series expansion of $f(x)$. Then, setting $x = 0$, we find that $a_0 = f(0)$. If we differentiate $\sum_{n=0}^{\infty} a_n x^n$ term by term, and again set $x = 0$, we find that $a_1 = f^{(1)}(0)$. Repeating this process, we find that, in general, $a_n = f^{(n)}(0)/n!$. From this we draw the important conclusion that if a function $f(x)$ is analytic in a neighborhood of the origin, then the nth coefficient in its power series expansion must be $f^{(n)}(0)/n!$, and this, among other things, proves the following result.

Theorem 3.9 *Suppose $f(x)$ is analytic in $(-r, r)$. Then the power series expansion of $f(x)$ is unique. That is, if $\sum_{n=0}^{\infty} a_n x^n$ and $\sum_{n=0}^{\infty} b_n x^n$ are two power series both of which converge to $f(x)$ in $(-r, r)$, then $a_n = b_n$, for $n = 0, 1, 2, \ldots$.*

Suppose now that $f(x)$ is C^{∞} in the interval $(-r, r)$. We wish to find out whether or not it is, in addition, analytic in $(-r, r)$. (NOTE: Any function that is a candidate for analyticity must be at least C^{∞}, since all analytic functions are C^{∞}. On the other hand, as we have pointed out before and as we shall shortly show, there are C^{∞} functions that are not analytic.) As we have noted, the nth coefficient of the power series expansion of $f(x)$, if it has one, must be $f^{(n)}(0)/n!$. Hence the nth partial sum $s_n(x)$ of the power series for $f(x)$, if it has one, must be

$$f(0) + f^{(1)}(0)x + \cdots + \frac{f^{(n)}(0)}{n!} x^n.$$

Remark We may ask why $\sum_{n=0}^{\infty} (f^{(n)}(0)/n!)x^n$ is not always a valid power series representation for *any* C^∞ function $f(x)$. The reason is, quite simply, that this series does not necessarily converge, except, of course, for $x = 0$, and moreover, if it does converge, it does not necessarily converge to $f(x)$.

Returning to the issue at hand, we see that the question is whether or not $f(x) - s_n(x) \to 0$ for each x in $(-r, r)$. In what follows, we assume that x is ≥ 0, since if x is negative, we can simply replace the function $f(t)$ by the function $f(-t)$, and consider the question of whether or not the power series for the latter converges for $t = |x|$. The expression (3.3), which we shall shortly obtain, is valid for x of any signature.

Now define $f_n(x) = f(x) - s_n(x)$. We then have to determine whether or not $f_n(x) \to 0$ for $x \in [0, r)$. We begin by noting that $f_n^{(j)}(0) = 0$, for $j = 0, 1, \ldots, n$, while for $j > n$, $f_n^{(j)}(x) = f^{(j)}(x)$ for all $x \in [0, r)$. From this it follows that for $x \in [0, r)$,

$$f_n(x) = \int_0^x f_n^{(1)}(t)\,dt$$

$$f_n^{(1)}(x) = \int_0^x f_n^{(2)}(t)\,dt$$

$$\vdots$$

$$f_n^{(n)}(x) = \int_0^x f_n^{(n+1)}(t)\,dt = \int_0^x f^{(n+1)}(t)\,dt.$$

Suppose now that $x \in [0, r)$ is fixed. Define $m_{n+1} = \min_{t\in[0,x]} f^{(n+1)}(t)$; $M_{n+1} = \max_{t\in[0,x]} f^{(n+1)}(t)$. Now $f_n^{(n)}(y) = \int_0^y f^{(n+1)}(t)\,dt$ for any $y \in [0, x]$, so we conclude that for $y \in [0, x]$,

$$m_{n+1}y \leq f_n^{(n)}(y) \leq M_{n+1}y.$$

Similarly, $f_n^{(n-1)}(y) = \int_0^y f_n^{(n)}(t)\,dt$, for $y \in [0, x]$, so

$$\int_0^y m_{n+1}t\,dt \leq f_n^{(n-1)}(y) \leq \int_0^y M_{n+1}t\,dt,$$

or

$$\tfrac{1}{2}m_{n+1}y^2 \leq f_n^{(n-1)}(y) \leq \tfrac{1}{2}M_{n+1}y^2,$$

for $y \in [0, x]$.

Continuing in this fashion, we ultimately find that

$$\frac{m_{n+1}}{(n+1)!}\,y^{n+1} \leq f_n(y) \leq \frac{M_{n+1}}{(n+1)!}\,y^{n+1},$$

for any $y \in [0, x]$, and in particular, therefore, for $y = x$. Thus, by the intermediate value theorem, there is a number $\theta \in [0, x]$, such that

$$f_n(x) = \frac{f^{(n+1)}(\theta)}{(n+1)!} x^{n+1}. \tag{3.3}$$

Finally, if we note that the preceding analysis only requires that the function f be of class C^{n+1}, we obtain the following theorem.

Theorem 3.10 (Taylor) *Suppose $f(x)$ is C^{n+1} in $(-r, r)$ and suppose $x \in (-r, r)$. Denote by I the closed interval with endpoints 0 and x. Then there is a number $\theta \in I$, such that*

$$f(x) - \left(f(0) + f^{(1)}(0)x + \cdots + \frac{f^{(n)}(0)}{n!} x^n\right) = \frac{f^{(n+1)}(\theta)}{(n+1)!} x^{n+1}.$$

Simple as it is, this is one of the most far-reaching theorems in analysis. Among other things, it provides a criterion for determining whether or not a function is analytic in a neighborhood of the origin, namely, it shows that $f(x)$ is certainly analytic in $(-r, r)$ if

$$\sup_{x \in (-r,r)} \frac{|f^{(n)}(x)| r^n}{n!} \to 0$$

as $n \to \infty$. Although it may appear complicated, this criterion is often very easy to apply in specific cases. We illustrate with two famous examples.

Examples

1. $f(x) = e^x$. In this case, $f^{(n)}(x) = e^x$, so we conclude that

$$e^x - \left(1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}\right) = \frac{e^\theta}{(n+1)!} x^{n+1},$$

where θ lies between 0 and x. In particular,

$$\left| e^x - \left(1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}\right)\right| \leq \frac{e^{\max(0,x)}|x|^{n+1}}{(n+1)!}.$$

Thus, since for fixed x,

$$\frac{e^{\max(0,x)}|x|^{n+1}}{(n+1)!} \to 0 \quad \text{as} \quad n \to \infty,$$

we conclude that the infinite series $1 + x + x^2/2! + x^3/3! + \cdots$ converges for all x, and, moreover, that its sum is e^x.

2. $f(x) = \sin x$. In this case, the successive derivatives of $f(x)$ are $\cos x$, $-\sin x$, $-\cos x$, $\sin x$, etc., so that in particular, $|f^{(n)}(x)| \leq 1$, for any combination of x and n. Thus by Taylor's theorem,

$$\left|\sin x - \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + (-1)^{(n-1)/2}\frac{x^n}{n!}\right)\right| \leq \frac{|x|^{n+1}}{(n+1)!}.$$

where n, of course, is odd. Now for fixed x, $|x|^{n+1}/(n+1)! \to 0$ as $n \to \infty$, so we conclude that the series $x - x^3/3! + x^5/5! - x^7/7! + \cdots$ converges for all x, and moreover, that the sum is $\sin x$.

Definition 3.2 If $f(x)$ is analytic in a neighborhood of the origin, we call the series $\sum_{n=0}^{\infty} (f^{(n)}(0)/n!)x^n$ the *Taylor series* for $f(x)$ or, sometimes, the Taylor series about the origin for $f(x)$.

4 A C^∞ function that is not analytic

We shall now show by an example that there exists a function that is C^∞ in a neighborhood of the origin (our example is, in fact, C^∞ on the whole line), but that is not analytic in any neighborhood of the origin. That is, the class of functions C^∞ in $(-r, r)$ is larger than the class of functions analytic in $(-r, r)$.

A function which serves as an example can be defined as follows: set $f(x) = e^{-1/x^2}$, for $x \neq 0$, and define $f(0) = 0$. Then it is evident that $f(x)$ is at least continuous (that is, of class C^0) on the whole line. In order to show that $f(x)$ is actually C^∞, we need the following fact: if $P(1/x)$ is any polynomial in $1/x$, then $P(1/x)e^{-1/x^2} \to 0$ as $x \to 0$ through nonzero values. To prove this, it suffices to show that for any fixed n, $(1/|x|)^n e^{-1/x^2} \to 0$ as $x \to 0$ through nonzero values, or equivalently that $x^n e^{-x^2} \to 0$ as $x \to \infty$. This, however, follows directly from l'Hospital's rule.

Suppose now it has been shown that $f(x)$ is C^n on the whole line and, moreover, that $f^{(n)}(0) = 0$, and $f^{(n)}(x) = P_n(1/x)e^{-1/x^2}$, for $x \neq 0$, where $P_n(1/x)$ is a polynomial in $1/x$. This is certainly true for $f^{(0)}(x)$, since in this case, we can simply take $P_0(1/x) = 1$. We want to show that $f(x)$ is C^{n+1}, and that $f^{(n+1)}(0) = 0$, while for $x \neq 0$, $f^{(n+1)}(x) = P_{n+1}(1/x)e^{-1/x^2}$, where $P_{n+1}(1/x)$ is a polynomial in $1/x$.

To show that $f^{(n)}(x)$ is differentiable at $x = 0$, we use the delta method. Consider

$$\frac{f^{(n)}(h) - f^{(n)}(0)}{h} = \frac{1}{h} P_n\left(\frac{1}{h}\right) e^{-1/h^2}.$$

By our Remark, this quantity tends to zero as $h \to 0$, which shows that $f^{(n)}(x)$ is differentiable at $x = 0$ and that the derivative is zero. For $x \neq 0$, it is clear from the form of $f^{(n)}(x)$ that $f^{(n)}(x)$ is continuously differentiable and that the derivative is of the form $P_{n+1}(1/x)e^{-1/x^2}$, where $P_{n+1}(1/x)$ is a polynomial in $1/x$. By our Remark, this tends to zero as $x \to 0$, and this proves that $f(x)$ is C^{n+1}, and moreover, that $f^{(n+1)}(0) = 0$, and $f^{(n+1)}(x) = P_{n+1}(1/x)e^{-1/x^2}$, if $x \neq 0$. By induction, this proves that $f(x)$ is actually C^∞ on the whole line, and, what is critical for our purposes, that $f^{(n)}(0) = 0$, for $n = 0, 1, 2, \ldots$.

The function $f(x)$, which we have just defined, furnishes the desired example: for suppose $f(x)$ were analytic in some neighborhood of the origin. Then each coefficient in the Taylor series for $f(x)$ would be 0, so that the series would converge to zero for all x. But $f(x) \neq 0$ unless $x = 0$. Thus $f(x)$ cannot be analytic in a neighborhood of the origin.

Remark By defining

$$\begin{aligned} f(x) &= e^{-1/(1-x^2)}, && \text{if } |x| < 1, \\ f(x) &= 0, && \text{if } |x| \geq 1, \end{aligned}$$

we obtain a function that is C^∞ on the whole line, positive for $|x| < 1$, and zero for $|x| \geq 1$. The proof is similar to the preceding discussion. Such functions are not only useful examples but are also very important tools in analysis.

5 Expansions about points other than the origin

Up to this point, we have, for reasons of clarity, only considered power series expansions about the origin. I.e., we have discussed power series in x, and the functions which possess such expansions. Frequently, however, the region of interest is not an interval about the origin, but an interval about some point $x_0 \neq 0$. Now the whole theory that we have built up can be trivially "translated" so as to encompass the more general questions that arise in this way. For example, suppose a function $f(x)$ is given in an interval of the form $|x - x_0| < r$ about the point x_0, and suppose we wish to determine whether or not $f(x)$ can be expanded in powers of $x - x_0$ within this interval. That is, we wish to know whether or not there exist numbers $a_0, a_1, a_2, \ldots$, such that the series $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ converges to $f(x)$ for $|x - x_0| < r$. Define $g(x) = f(x + x_0)$. Then $g(x)$ is defined in the interval $|x| < r$, and it is clear that $f(x)$ can be expressed as a power series in $x - x_0$, in the interval $|x - x_0| < r$, if and only if $g(x)$ can be expressed as a power

series in x, in the interval $|x| < r$; that is, if and only if $\sum_{n=0}^{\infty} (g^{(n)}(0)x^n/n!)$ converges to $g(x)$ for $|x| < r$. Now $g^{(n)}(0) = f^{(n)}(x_0)$, so the question becomes whether or not $\sum_{n=0}^{\infty} (f^{(n)}(x_0)x^n/n!)$ converges to $g(x)$ for $|x| < r$, or, equivalently, whether or not $\sum_{n=0}^{\infty} (f^{(n)}(x_0)(x - x_0)^n/n!)$ converges to $f(x)$ for $|x - x_0| < r$. The last series is called the Taylor expansion of $f(x)$ about the point $x = x_0$.

As we have pointed out, our entire theory can be "translated" in the above way. For instance, to every series of the form $\sum_{n=0}^{\infty} a_n(x - x_0)^n$, there corresponds a radius of convergence r, having the property that the series is absolutely convergent if x satisfies $|x - x_0| < r$, and diverges if $|x - x_0| > r$. If $r = 0$, the series converges only for $x = x_0$, while if $r = \infty$, the series is absolutely convergent for all x. It is trivial to show that the methods for determining the radius of convergence (for example, the ratio method, the root method, or, most generally, the Cauchy-Hadamard method) carry over verbatim to the more general context.

6 A special result

We shall require the following special result in the next chapter.

Theorem 3.11 *For any $\epsilon > 0$, there exists a sequence $\{P_n(x)\}$ of polynomials that converges uniformly to the function $(x + \epsilon)^{1/2}$—the positive square root of $x + \epsilon$, on $[0,1]$.*

PROOF Suppose $\epsilon > 0$ is given. By translation, the assertion is equivalent to the statement that there exists a sequence $\{P_n^*(x)\}$ of polynomials that converges uniformly to $(x + \frac{1}{2} + \epsilon)^{1/2}$ on $[-\frac{1}{2}, \frac{1}{2}]$; for then the sequence of polynomials $\{P_n^*(x - \frac{1}{2})\}$ converges uniformly to $(x + \epsilon)^{1/2}$ on $[0, 1]$. Now $(x + \frac{1}{2} + \epsilon)^{1/2} = c^{1/2}(1 + x/c)^{1/2}$, where $c = \frac{1}{2} + \epsilon$, so it clearly suffices to show that there exists a sequence of polynomials that converges uniformly to $(1 + x/c)^{1/2}$ on $[-\frac{1}{2}, \frac{1}{2}]$. In order to show this, we shall, in fact, show that $(1 + x/c)^{1/2}$ is analytic in the interval $(-c, c)$, which implies that the partial sums of the power series for $(1 + x/c)^{1/2}$ provide an appropriate sequence of polynomials, since $c > \frac{1}{2}$. In order to show that $(1 + x/c)^{1/2}$ is representable by a power series in $(-c, c)$, we consider the binomial series

$$\sum_{n=0}^{\infty} a_n x^n, \tag{3.4}$$

where $a_0 = 1$ and $a_n = \dfrac{\frac{1}{2}(\frac{1}{2} - 1)(\frac{1}{2} - 2)\cdots(\frac{1}{2} - (n - 1))}{c^n n!}$ for $n \geq 1$.

Note that it follows from a trivial calculation that for $n = 0, 1, 2, \ldots$ $c(n + 1)a_{n+1} = (\frac{1}{2} - n)a_n$, or, equivalently, $c(n + 1)a_{n+1} + na_n = \frac{1}{2}a_n$.

Now $|a_{n+1}/a_n| = (1/c)|(n - \frac{1}{2})/(n + 1)|$, so it follows from the ratio test that (3.4) has radius of convergence c. It remains to show that (3.4) converges to $(1 + x/c)^{1/2}$.

Denote by $f(x)$ the function to which (3.4) converges in $(-c, c)$, and consider the quotient $f(x)(1 + x/c)^{-1/2}$. We show that the derivative of the quotient is identically zero in $(-c, c)$. This implies that the quotient must be a constant. Moreover, the value of the constant must be 1, since the value of the quotient at $x = 0$ is 1. This implies that $f(x) = (1 + x/c)^{1/2}$ for $x \in (-c, c)$.

It remains to show that

$$\left(f(x)\left(1 + \frac{x}{c}\right)^{-1/2}\right)' = 0, \qquad \text{for } x \in (-c, c).$$

Now

$$\left(f(x)\left(1 + \frac{x}{c}\right)^{-1/2}\right)' = f'(x)\left(1 + \frac{x}{c}\right)^{-1/2} - \frac{1}{2c}f(x)\left(1 + \frac{x}{c}\right)^{-3/2}$$

Multiplying the last expression by $c(1 + x/c)^{-3/2}$, we see that to show that $(f(x)(1 + x/c)^{-1/2})'$ vanishes identically in $(-c, c)$, it suffices to show that $f'(x)(c + x) = \frac{1}{2}f(x)$, for $x \in (-c, c)$.

Now

$$cf'(x) = \sum_{n=1}^{\infty} cna_nx^{n-1}$$

$$= \sum_{n=0}^{\infty} c(n + 1)a_{n+1}x^n.$$

Similarly,

$$xf'(x) = \sum_{n=1}^{\infty} na_nx^n$$

$$= \sum_{n=0}^{\infty} na_nx^n.$$

That is,

$$f'(x)(c + x) = \sum_{n=0}^{\infty} (c(n + 1)a_{n+1} + na_n)x^n$$

$$= \sum_{n=0}^{\infty} \tfrac{1}{2}a_nx^n$$

$$= \tfrac{1}{2}f(x),$$

and this completes the proof.

7 Concluding remarks

Some aspects of power series become much clearer if considered in terms of the theory of functions of a complex variable. This is particularly true of the material in Section 4. For example, even leaving aside Theorem 3.2, it is perfectly clear why the power series $1 + x + x^2 + x^3 + \cdots$, which represents $1/(1-x)$, has radius of convergence 1. In this case, the function $1/(1-x)$ itself tends to infinity as $x \to 1$. On the other hand, there is no such obvious reason to explain why the series $1 - x^2 + x^4 - x^6 + \cdots$, which represents $1/(1+x^2)$, should also have radius of convergence 1, since the function $1/(1+x^2)$ is perfectly well-behaved over the entire real line. Without going deeply into this question, which we cannot do without the apparatus of complex variables, it is worthwhile to point out that the function $1/(1+x^2)$ does possess "singularities" at the points i and $-i$ of the complex plane, both of which are at distance 1 from the origin, and it is this fact that accounts, from the above point of view, for the radius of convergence of $1 - x^2 + x^4 - x^6 + \cdots$ being 1. For an excellent introduction to this and other topics, the interested reader should consult [3].

Exercises

1 We have already seen that $x - (x^2/2) + (x^3/3) - (x^4/4) + \cdots$ converges to $\log(1+x)$ on $(-1, 1)$. Using the fact that if $\{a_n\}$ is a positive sequence that decreases to zero, then $|\sum_{n=N}^{\infty}(-1)^n a_n| \le a_N$, show that the series for $\log(1+x)$ converges uniformly on $[0, 1]$, and hence that $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = \log 2$.

2 Suppose $\sum_{n=0}^{\infty} a_n$ converges. Show that $\lim_{x\uparrow 1} \sum_{n=0}^{\infty} a_n x^n$ exists and equals the sum of $\sum_{n=0}^{\infty} a_n$. This provides an alternate derivation of the result in Problem 1. (HINT: set $A_n = a_0 + \cdots + a_n$. Then

$$\sum_{n=0}^{\infty} a_n x^n = (1-x)\sum_{n=0}^{\infty} A_n x^n.$$

Moreover, $A_n = c - e_n$, where c is the sum of $\sum_{n=0}^{\infty} a_n$, and $e_n \to 0$. Thus, $c - \sum_{n=0}^{\infty} a_n x^n = c - (1-x)\sum_{n=0}^{\infty}(c - e_n)x^n = (1-x)\sum_{n=0}^{\infty} e_n x^n$, so the problem becomes one of showing that the last quantity tends to zero as $x \uparrow 1$.)

3 Suppose $f(x) = \sum_{n=0}^{\infty} x^{n^2}$, for x in $(-1, 1)$. Show that $[f(x)]^2 = \sum_{n=0}^{\infty} r(n)x^n$, where $r(n)$ is the number of representations of n an ordered sum of two non-negative squares.

4 Suppose $f(x)$ is analytic in $(-r, r)$, and suppose $\{x_n\}$ is a sequence of points in $(-r, r)$ such that $x_n \to 0$, and $f(x_n) = 0$, for all n. Show that $f(x) \equiv 0$ in $(-r, r)$. (HINT: suppose $\sum_{n=0}^{\infty} a_n x^n$ is the power series for $f(x)$. It follows from the continuity of $f(x)$ that $a_0 = 0$. Prove by induction that $a_n = 0$ for $n = 1, 2, 3, \ldots$.)

5 Find the power series about the origin for $\tan^{-1} x$, by considering

$$\int_0^x (1 + t^2)^{-1}\, dt.$$

6 Find a solution to the differential equation $y'' + y = x$, $y(0) = 0$, $y'(0) = 1$, by assuming that y has a power series expansion, and then solving for the coefficients.

7 Show that e is irrational. (HINT: suppose $p/q = 1 + 1/1! + 1/2! + \cdots$. Then $q!(1 + 1/1! + 1/2! + \cdots)$ would be an integer, and hence so would

$$\frac{1}{(q+1)} + \frac{1}{(q+1)(q+2)} + \frac{1}{(q+1)(q+2)(q+3)} + \cdots.$$

Show that this is impossible, since the last quantity is less than 1.)

8 Show that $\lim_{n\to\infty} (1 + 1/n)^n$ exists and equals e. (HINT: it suffices to show that $n \log (1 + 1/n) \to 1$. Use the power series for $\log (1 + x)$.)

9 In general, is $(1 + 1/n)^n$ smaller or larger than e?

10 Suppose i is a quantity satisfying $i^2 = -1$. On a purely formal basis, derive the result $e^{ix} = \cos x + i \sin x$.

CHAPTER FOUR

The Weierstrass Approximation Theorem

1 The Weierstrass approximation theorem

Suppose $f(x)$ is analytic in an interval $(-r, r)$, and suppose $[-r', r']$ is a closed subinterval of $(-r, r)$. Then the Taylor series for $f(x)$ about $x = 0$ converges uniformly to $f(x)$ on $[-r', r']$. In particular, therefore, there exists a sequence of polynomials that converges uniformly to $f(x)$ on $[-r', r']$, namely, the partial sums of the Taylor series.

Now as we have seen, not all functions, not even all C^{∞} functions, are analytic on $(-r, r)$, so we cannot in general expect that there is a power series whose partial sums converge uniformly to a given function on subintervals of $(-r, r)$. Nevertheless, we may pose the following question: if $[a, b]$ is a closed interval, for what functions $f(x)$ on $[a, b]$ does a sequence of polynomials exist that converges uniformly to $f(x)$ on $[a, b]$? Or equivalently, for what functions $f(x)$ is it true that for every $\epsilon > 0$, there exists a polynomial $P(x)$, such that $|f(x) - P(x)| < \epsilon$ for $x \in [a, b]$, that is, what functions $f(x)$ can be uniformly approximated by polynomials on $[a, b]$?

Remark 1 It should be borne in mind that there is no requirement for the approximating polynomials to be the partial sums of some fixed power series. So the class of functions that can be uniformly approximated by polynomials over $[a, b]$ is, if anything, larger than the class of functions that are analytic in an interval containing $[a, b]$.

Remark 2 It is evident that any function that can be uniformly approximated on $[a, b]$ by polynomials, which are automatically continuous, must itself be continuous (Theorem 2.8).

The remarkable theorem, due to Weierstrass, that settles this whole question is the following.

Theorem 4.1 **(Weierstrass Approximation Theorem)** *Suppose $[a, b]$ is a closed interval, and suppose $f(x)$ is continuous on $[a, b]$. Then $f(x)$ can be uniformly approximated by polynomials on $[a, b]$. For any $\epsilon > 0$, there exists a polynomial $P(x)$, such that $|f(x) - P(x)| < \epsilon$ for $x \in [a, b]$, or equivalently, there exists a sequence $\{P_n(x)\}$ of polynomials, such that $P_n(x) \to f(x)$ uniformly on $[a, b]$.*

PROOF We begin with a series of successive simplifications, by which we shall ultimately show that it suffices to prove the theorem for the single case in which $f(x) = |x|$, and $[a, b] = [-1, 1]$.

The first simplification consists in showing that it suffices to prove the theorem for piecewise linear continuous functions, or, for short, polygonal functions, where by a polygonal function on $[a, b]$, we mean a continuous function whose graph consists of a finite number of segments of straight lines (Figure 8). We call the points on the graph at which the slope changes, the "vertices" of the graph, and as a matter of convention, we include the points $(a, f(a))$, $(b, f(b))$ among the vertices.

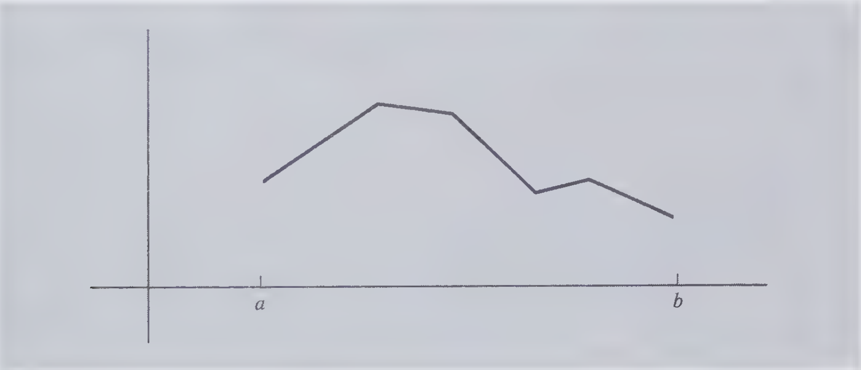

FIGURE 8

Now any continuous function $f(x)$ on $[a, b]$ can be uniformly approximated by polygonal functions. This follows almost immediately from the uniform continuity of $f(x)$ on $[a, b]$. For suppose we divide the interval $[a, b]$ into n equal subintervals, by introducing the points of subdivision $x_0, x_1, \ldots, x_n$, where $x_0 = a$, $x_n = b$. Define $g_n(x)$ to be the polygonal function that is linear on each subinterval, and whose value at x_k is $f(x_k)$. (See Figure 9, where the case $n = 4$ is illustrated.)

Now as the number of points of subdivision becomes large, over any subinterval the maximum oscillation of $f(x)$, hence of $g_n(x)$, and hence of $f(x) - g_n(x)$, tends to zero, which implies that $g_n(x) \to f(x)$ uniformly over $[a, b]$, since $f(x) - g_n(x) = 0$ at each point of subdivision.

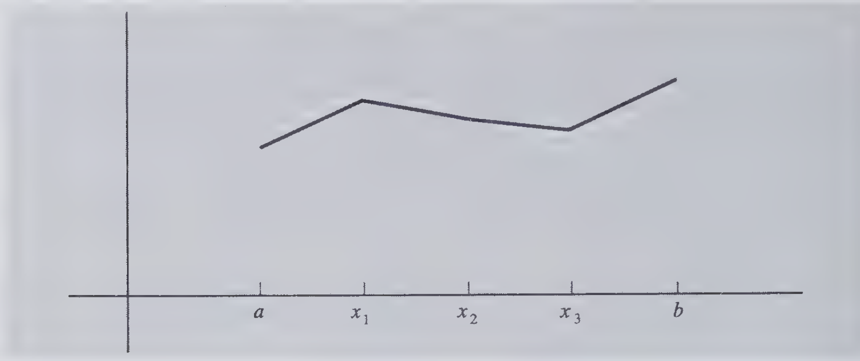

FIGURE 9

Suppose now that the theorem has been established for polygonal functions. I.e., suppose that if $g(x)$ is a polygonal function, then for any $\epsilon > 0$, there exists a polynomial $P(x)$, such that $|g(x) - P(x)| < \epsilon$ for $x \in [a, b]$. To show that this implies the theorem for any continuous function, suppose $f(x)$ is continuous on $[a, b]$, and suppose $\epsilon > 0$ is given. By what has already been said, we know that there exists a polygonal function $g(x)$, such that $|f(x) - g(x)| < \frac{1}{2}\epsilon$ for $x \in [a, b]$. Now if the Weierstrass theorem is true for $g(x)$, there exists a polynomial $P(x)$, such that $|g(x) - P(x)| < \frac{1}{2}\epsilon$ for $x \in [a, b]$. Thus, by the triangle inequality, $|f(x) - P(x)| < \epsilon$ for $x \in [a, b]$. Summarizing, we have shown that to obtain the general result, it suffices to prove the theorem for polygonal functions.

To obtain the next simplification of the problem, we note that on $[a, b]$, any polygonal function can be expressed as the sum of a constant and a linear combination of translates of the function $|x|$. More precisely, if $f(x)$ is a polygonal function on $[a, b]$, whose graph has vertices, from left to right, at the points $(v_0, w_0), \ldots, (v_n, w_n)$, then there exist numbers $a_0, \ldots, a_{n-1}, a_n$, such that

$$f(x) = a_0|x - v_0| + \cdots + a_{n-1}|x - v_{n-1}| + a_n.$$

To prove this, we note that a polygonal function $f(x)$ on $[a, b]$ is completely determined by specifying $f(a)$, the x coordinates of the vertices, and the slopes of the segments connecting them. Suppose the slopes of the successive segments are $s_0, \ldots, s_{n-1}$. Then the existence of a representation of the desired kind is equivalent to the existence of numbers $a_0, \ldots, a_{n-1}$, satisfying

$$\begin{aligned} a_0 - a_1 - a_2 - \cdots - a_{n-1} &= s_0 \\ a_0 + a_1 - a_2 - \cdots - a_{n-1} &= s_1 \\ &\vdots \\ a_0 + a_1 + a_2 + \cdots + a_{n-1} &= s_{n-1} \end{aligned} \tag{4.1}$$

These equations simply express the fact that the contribution to the slope of the ith segment from the function $a_j|x - v_j|$ is either a_j or $-a_j$, depending on whether or not the point v_j lies to the left or right of the initial point of the ith segment. We solve this system as follows: adding the first equation to the last, we obtain $a_0 = \frac{1}{2}(s_0 + s_{n-1})$. Discarding the first equation, and subtracting the above value of a_0 from all the remaining equations, we end up with the system

$$\begin{aligned} a_1 - a_2 - a_3 - \cdots - a_{n-1} &= s_1^* \\ a_1 + a_2 - a_3 - \cdots - a_{n-1} &= s_2^* \\ &\vdots \\ a_1 + a_2 + a_3 + \cdots + a_{n-1} &= s_{n-1}^* \end{aligned}$$

where $s_j^* = s_j - \frac{1}{2}(s_0 + s_{n-1})$. By repeating the above process, we solve successively for $a_1, \ldots, a_{n-1}$, and ultimately end up with the complete solution of the system (4.1).

We have now reduced the problem of proving Weierstrass' theorem to the problem of proving it for a function of the form $|x - v_j|$ on $[a, b]$. For if $P_1^j(x), P_2^j(x), \ldots$ is a sequence of polynomials that converges uniformly to $|x - v_j|$ on $[a, b]$, it is clear that the sequence of polynomials

$$\{a_0 P_k^1(x) + \cdots + a_{n-1} P_k^{n-1}(x) + a_n\}_{k=1}^{\infty}$$

converges uniformly to $a_0|x - v_0| + \cdots + a_{n-1}|x - v_{n-1}| + a_n$ on $[a, b]$.

To show that there exists a sequence $\{P_k^j(x)\}_{k=1}^{\infty}$ of polynomials converging uniformly to $|x - v_j|$ on $[a, b]$, it suffices to prove that there exists a sequence $\{Q_k(x)\}_{k=1}^{\infty}$ of polynomials converging uniformly to $|x|$ on $[-(b - a), (b - a)]$, since it is then clear that $\{Q_k(x - v_j)\}_{k=1}^{\infty}$ has the desired properties. That is, Weierstrass' theorem will be proved if we can show that for any $r > 0$, there exists a sequence $\{Q_k(x)\}_{k=1}^{\infty}$ of polynomials such that $Q_k(x) \to |x|$ uniformly on $[-r, r]$. In doing this, we may assume $r = 1$; for if $Q_k(x) \to |x|$ uniformly on $[-1, 1]$, it is clear that $rQ_k(x/r) \to |x|$ uniformly on $[-r, r]$, for fixed r. I.e., the whole question comes down to the following: given $\epsilon > 0$, does there exist a polynomial $P(x)$, such that $||x| - P(x)| < \epsilon$ on $[-1, 1]$?

To answer this in the affirmative, suppose $y > 0$ is small enough so that $(x^2 + y)^{1/2} - |x|$ is less than $\frac{1}{2}\epsilon$ on $[-1, 1]$. Now by the concluding remarks of the last chapter, there exists a polynomial $Q(x)$, such that

$$|(x + y)^{1/2} - Q(x)| < \tfrac{1}{2}\epsilon$$

for $x \in [0, 1]$. Thus clearly $|(x^2 + y)^{1/2} - Q(x^2)| < \frac{1}{2}\epsilon$ for $x \in [-1, 1]$, and by the triangle inequality, this implies that $||x| - Q(x^2)| < \epsilon$ for $x \in [-1, 1]$. Therefore, if we take $P(x) = Q(x^2)$, Weierstrass' theorem is proved.

Remark 1 The Weierstrass theorem is not necessarily true for continuous functions on *open* intervals. This is shown by the function $f(x) = 1/x$, which is continuous on $(0, 1)$, but cannot be uniformly approximated by a polynomial on $(0, 1)$, since it is not even bounded.

Remark 2 The following simple consequence of the Weierstrass theorem is worth pointing out. Suppose $f(x)$ is a continuous function on $[a, b]$ that is strictly increasing: $f(x_1) < f(x_2)$ for $a \leq x_1 < x_2 \leq b$. Then $f(x)$ has a continuous inverse $f^{-1}(x)$, defined on the interval $[a', b']$, where $a' = f(a)$; $b' = f(b)$. By this, we mean, as usual, that $f^{-1}(f(x)) = x$ for $x \in [a, b]$, and $f(f^{-1}(x)) = x$ for $x \in [a', b']$. Now we can set up a 1–1 correspondence between the set of continuous functions on $[a, b]$, and the set of continuous functions on $[a', b']$, as follows. To each continuous function $g(x)$ on $[a, b]$, there corresponds the continuous function $g^*(x) = g(f^{-1}(x))$ on $[a', b']$. Similarly, to each continuous function $g^*(x)$ on $[a', b']$, there corresponds the continuous function $g(x) = g^*(f(x))$ on $[a, b]$. Now it is a simple matter to verify that if $g_n^*(x) \to g^*(x)$ uniformly on $[a', b']$, then $g_n(x) \to g(x)$ uniformly on $[a, b]$, and vice versa. Moreover, it is clear that the function on $[a', b']$ that corresponds in the above way to $f(x)$ is simply the identity, or in the usual notation, x. Suppose now that $g(x)$ is a continuous function on $[a, b]$. Then we know by the Weierstrass theorem that the corresponding function $g^*(x)$ on $[a', b']$ can be uniformly approximated by polynomials. Accordingly, let $P_n(x) \to g^*(x)$ uniformly on $[a', b']$. Now the function on $[a, b]$ that corresponds to $P_n(x)$ is simply $P_n(f(x))$, so we conclude that $P_n(f(x)) \to g(x)$ uniformly on $[a, b]$. That is, we have shown that if $f(x)$ is continuous and strictly increasing on $[a, b]$, then any continuous function on $[a, b]$ can be uniformly approximated by polynomials in $f(x)$. That is, by expressions of the form

$$a_0 + a_1 f(x) + a_2(f(x))^2 + \cdots + a_n(f(x))^n.$$

(Note that the argument applies equally well to strictly decreasing functions $f(x)$, if we substitute $-f(x)$ for $f(x)$.)

Example

Consider the function $\cos x$. This function is strictly increasing on $[-\pi, 0]$, and hence, by the above, any continuous function on $[-\pi, 0]$ can be uniformly approximated by polynomials in $\cos x$. Now $\cos x$ is an *even* function on $[-\pi, \pi]$, that is, $\cos -x = \cos x$. From this, it is easy to see that any even function on $[-\pi, \pi]$ can be uniformly approximated by polynomials in $\cos x$. Finally, if we note that

$$(\cos Ax)(\cos Bx) = \tfrac{1}{2}[\cos (A + B)x + \cos (A - B)x],$$

it follows by an obvious induction that $\cos^n x$ can be expressed as a linear combination of the functions 1, $\cos x$, $\cos 2x, \ldots, \cos nx$, and this in turn implies that any continuous even function on $[-\pi, \pi]$ can be uniformly approximated by linear combinations of the functions 1, $\cos x$, $\cos 2x$, $\cos 3x, \ldots$.

Remark 3 There is a very beautiful generalization of the Weierstrass theorem, due to M. H. Stone, which, in out context, goes as follows.

Theorem 4.2 (Stone-Weierstrass) *Suppose F is a collection of continuous functions on an interval* $[a, b]$, *and suppose the collection F has the following properties*:

1. *If* $f(x)$ *is in F, then so is* $cf(x)$, *for any constant c.*
2. *If* $f(x)$ *and* $g(x)$ *are in F, then so are* $f(x) + g(x)$ *and* $f(x)g(x)$.
3. *For any two distinct points* $x_0, x_1 \in [a, b]$, *there exists a function* $f(x)$ *in F, such that* $f(x_0) \neq f(x_1)$.
4. *For any* $x_0 \in [a, b]$, *there exists a function* $f(x)$ *in F, such that* $f(x_0) \neq 0$.

Then any continuous function on $[a, b]$ *can be uniformly approximated by functions in F. That is, if* $g(x)$ *is continuous on* $[a, b]$, *and if* $\epsilon > 0$ *is given, then there exists a function* $f(x)$ *in F, such that* $|g(x) - f(x)| < \epsilon$ *for* $x \in [a, b]$.

We shall not give a proof of this theorem, but the interested reader can find a very good account in [5].

Remark 4 The following result is an interesting example of the Weierstrass theorem. Suppose $f(x)$ is continuous on $[a, b]$, and suppose, moreover, that

$$\int_a^b x^n f(x)\, dx = 0, \qquad \text{for } n = 0, 1, 2, \ldots \tag{4.2}$$

(we define x^0 to be identically 1); then $f(x)$ is identically zero.

Proof: By the Weierstrass theorem, there exists a sequence $\{P_n(x)\}$ of polynomials such that $P_n(x) \to f(x)$ uniformly on $[a, b]$. Thus

$$\int_a^b (f(x) - P_n(x))^2\, dx \to 0 \qquad \text{as } n \to \infty.$$

I.e.,

$$\int_a^b (f(x))^2\, dx + \int_a^b (P_n(x))^2\, dx - 2\int_a^b f(x)P_n(x)\, dx \to 0$$

as $n \to \infty$. But

$$2\int_a^b f(x)P_n(x)\,dx = 0, \qquad \text{for any } n,$$

by (4.2). We conclude that

$$\int_a^b (f(x))^2\,dx + \int_a^b (P_n(x))^2\,dx \to 0 \qquad \text{as } n \to \infty,$$

and since both integrals are ≥ 0, we infer that $\int_a^b (f(x))^2\,dx = 0$. This implies that $f(x) \equiv 0$, for if not, there is, by continuity, a subinterval $[a', b']$ of $[a, b]$ on which $(f(x))^2 \geq c$, for some positive constant c. This would imply that $\int_a^b (f(x))^2\,dx \geq (b' - a')c$, which is impossible.

2 The Karamata Tauberian theorem

A striking application of the Weierstrass theorem is given by the Karamata Tauberian theorem. Before we give the precise statement of this theorem, it is perhaps worthwhile to discuss a concrete case, and fortunately, a simple example is provided by the geometric series. Accordingly, consider the series $1 + x + x^2 + \cdots$, which, as we know, converges to $1/(1 - x)$ for $|x| < 1$. Now as $x \to 1$, the value of the series becomes very large. More precisely, it becomes large in such a way that the product

$$(1 - x)\left(\sum_{n=0}^{\infty} x^n\right) \tag{4.3}$$

tends to 1 as $x \to 1$. (This result is, of course, trivial since the value of the product is precisely 1 for $|x| < 1$.)

Now besides the relationship (4.3), which describes in a fairly precise way how the function represented by the series tends to infinity as x tends to 1, we might also note the following obvious feature of the geometric series: the average value of the first n coefficients tends to 1 as $n \to \infty$. That is, if s_{n-1} denotes the sum of the first n coefficients, then $s_{n-1}/n \to 1$ as $n \to \infty$. (This is again trivial, since in this particular case, $s_{n-1} = n$ for $n = 1, 2, \ldots$.)

Now it is a very remarkable and important fact that the analogue of this relationship holds true for a very wide class of series. More precisely, we shall show, and this is Karamata's theorem, that if $\sum_{n=0}^{\infty} a_n x^n$ is any series with non-negative coefficients that converges in $|x| < 1$, and if, moreover, for some constant c,

$$(1 - x)\sum_{n=0}^{\infty} a_n x^n \to c \qquad \text{as } x \to 1, \tag{4.4}$$

then

$$\frac{1}{n+1}\sum_{j\le n} a_j \to c \qquad \text{as } n \to \infty.$$

In other words, information, of the type contained in (4.4), about the function to which a series converges, carries with it, in many cases, considerable information about what we might call the "average" size of the coefficients.

The proof of this theorem depends on an ingenious observation, which again we illustrate for the simple case of the geometric series. Suppose $f(x)$ is a Riemann integrable function on $[0, 1]$, and consider the series

$$(1 - x)\sum_{n=0}^{\infty} x^n f(x^n), \tag{4.5}$$

which clearly converges for $x \in [0, 1)$, since $f(x)$ is bounded on $[0, 1]$. (Throughout the following, we shall be dealing only with a very special subclass of Riemann integrable functions, namely, the functions that are continuous on $[0, 1]$, together with the function $h(x)$, defined as follows:

$$h(x) = 0, \qquad \text{for } x \in \left[0, \frac{1}{e}\right),$$

$$h(x) = \frac{1}{x}, \qquad \text{for } x \in \left[\frac{1}{e}, 1\right].$$

The importance of the function $h(x)$ will become apparent later on.)

Suppose now that x is in $(0, 1)$. We wish to take a closer look at (4.5). Now the numbers $1, x, x^2, x^3, \ldots$ divide the interval $(0, 1]$ into an infinite collection of subintervals $[x, 1], [x^2, x], [x^3, x^2], \ldots$. The length of the subinterval $[x^{n+1}, x^n]$ is $x^n(1 - x)$, which clearly tends to zero uniformly in n, as $x \to 1$. Moreover, since x^n is the right endpoint of $[x^{n+1}, x^n]$, it is not at all unreasonable to expect that as $x \to 1$, $(1 - x)\sum_{n=0}^{\infty} x^n f(x^n)$ tends to $\int_0^1 f(t)\,dt$, and this is, in fact, always the case. (Figure 10 illustrates what is happening in the special case $f(t) = t$; $x = \frac{1}{2}$. The sum $(1 - x)\sum_{n=0}^{\infty} x^n f(x^n)$ is precisely the total area of all the rectangles.)

Now it turns out that this situation holds for a wide class of series. More precisely, if $\sum_{n=0}^{\infty} a_n x^n$ is a series that converges for $|x| < 1$, with each $a_n \ge 0$, and if, moreover, $(1 - x)\sum_{n=0}^{\infty} a_n x^n \to c$ as $x \to 1$, and if, finally, $f(x)$ is Riemann integrable, then $(1 - x)\sum_{n=0}^{\infty} a_n x^n f(x^n) \to c\int_0^1 f(t)\,dt$, as $x \to 1$.

Curiously, this last fact is sufficient to prove Karamata's theorem. To show this, we return to the function $h(x)$, which we defined before. Note first of all that

$$\int_0^1 h(t)\,dt = \int_{1/e}^1 \frac{1}{t}\,dt = 1.$$

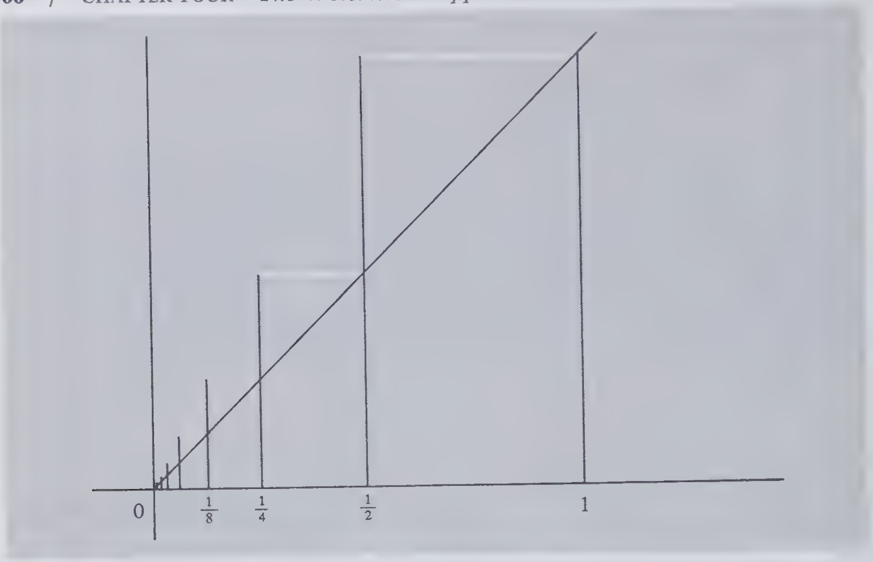

FIGURE 10

Suppose, now, that we have succeeded in proving that

$$(1 - x) \sum_{n=0}^{\infty} a_n x^n h(x^n) \to c \int_0^1 h(t)\, dt = c$$

as $x \to 1$.

Now $h(t) = 0$ for $t \in [0, 1/e)$, and $h(t) = 1/t$ for $t \in [1/e, 1]$, so

$$\begin{aligned}(1 - x) \sum_{n=0}^{\infty} a_n x^n h(x^n) &= (1 - x) \sum_{n \le -(1/\log x)} a_n x^n \frac{1}{x^n} \\ &= (1 - x) \sum_{n \le -(1/\log x)} a_n,\end{aligned}$$

where the last two sums are taken, as indicated, over all integers n for which $n \le -(1/\log x)$, or what is the same thing, for which $x^n \ge 1/e$.

For convenience, we set $y = -(1/\log x)$. Then $y \to \infty$ as $x \to 1$. Also,

$$(1 - x) \sum_{n \le -(1/\log x)} a_n = (1 - e^{-1/y}) \sum_{n \le y} a_n,$$

and by our temporary hypothesis, the last quantity tends to c as $y \to \infty$.

Now by taking the power series for $e^{-1/y}$, we see that

$$\frac{1 - e^{-1/y}}{1/(y + 1)} \to 1 \qquad \text{as } y \to \infty,$$

and this implies that

$$\frac{1}{y+1}\sum_{n\le y} a_n \to c \qquad \text{as } y \to \infty,$$

which implies Karamata's theorem.

Summarizing, we have shown that to prove Karamata's theorem, it suffices to show that if each $a_n \ge 0$ and if $(1-x)\sum_{n=0}^{\infty} a_n x^n \to c$ as $x \to 1$, then $(1-x)\sum_{n=0}^{\infty} a_n x^n h(x^n) \to c$ as $x \to 1$.

Accordingly, suppose $\sum_{n=0}^{\infty} a_n x^n$ is a series that converges for $|x| < 1$, and for which each $a_n \ge 0$. Suppose, moreover, that $(1-x)\sum_{n=0}^{\infty} a_n x^n \to c$ as $x \to 1$. We begin by showing that if $f(x)$ is a polynomial, then

$$(1-x)\sum_{n=0}^{\infty} a_n x^n f(x^n) \to c\int_0^1 f(t)\,dt \qquad \text{as } x \to 1.$$

In order to show this, it clearly suffices to consider the case in which $f(x) = x^k$, for some integer $k \ge 0$. In this case, $f(x^n) = x^{kn}$, so that

$$\begin{aligned}
&(1-x)\sum_{n=0}^{\infty} a_n x^n f(x^n)\\
&\quad = (1-x)\sum_{n=0}^{\infty} a_n x^n x^{kn} = (1-x)\sum_{n=0}^{\infty} a_n (x^{k+1})^n\\
&\quad = \frac{1-x^{k+1}}{1+x+\cdots+x^k}\sum_{n=0}^{\infty} a_n (x^{k+1})^n = \frac{1-y}{1+x+\cdots+x^k}\sum_{n=0}^{\infty} a_n y^n,
\end{aligned}$$

where we have set $y = x^{k+1}$. Now $y \to 1$ as $x \to 1$, so it follows from our hypotheses that the last quantity tends to

$$\frac{c}{k+1} = c\int_0^1 t^k\,dt,$$

which was to be shown.

Suppose now that $g(x)$ is a Riemann integrable function on $[0, 1]$, having the property that for any $\epsilon > 0$, there exist Riemann integrable functions $m(x)$ and $M(x)$, such that

1. $m(x) \le g(x) \le M(x)$ for all $x \in [0, 1]$

2a. $c\int_0^1 g(t)\,dt - \epsilon < c\int_0^1 m(t)\,dt$

b. $c\int_0^1 M(t)\,dt < \epsilon + c\int_0^1 g(t)\,dt$

3a. $(1 - x) \sum_{n=0}^{\infty} a_n x^n m(x^n) \to c \int_0^1 m(t)\, dt \qquad \text{as } x \to 1$

b. $(1 - x) \sum_{n=0}^{\infty} a_n x^n M(x^n) \to c \int_0^1 M(t)\, dt \qquad \text{as } x \to 1$

then

$$(1 - x) \sum_{n=0}^{\infty} a_n x^n g(x^n) \to c \int_0^1 g(t)\, dt \qquad \text{as } x \to 1 \tag{4.6}$$

To see this, suppose that, for a given $\epsilon > 0$, $m(x)$ and $M(x)$ have properties (1)–(3) above. Then, since $a_n \geq 0$ for each n, it follows that for any $x \in [0, 1)$,

$$\begin{aligned} (1 - x) \sum_{n=0}^{\infty} a_n x^n m(x^n) &\leq (1 - x) \sum_{n=0}^{\infty} a_n x^n g(x^n) \\ &\leq (1 - x) \sum_{n=0}^{\infty} a_n x^n M(x^n). \end{aligned} \tag{4.7}$$

Now it follows from (4.7) and from requirements (1)–(3) above, that there exists some $x_0 \in (0, 1)$, such that for all $x \in [x_0, 1)$,

$$c \int_0^1 g(t)\, dt - 2\epsilon \leq (1 - x) \sum_{n=0}^{\infty} a_n x^n g(x^n) \leq c \int_0^1 g(t)\, dt + 2\epsilon$$

That is,

$$-2\epsilon < (1 - x) \sum_{n=0}^{\infty} a_n x^n g(x^n) - c \int_0^1 g(t)\, dt < 2\epsilon,$$

or

$$\left| (1 - x) \sum_{n=0}^{\infty} a_n x^n g(x^n) - c \int_0^1 g(t)\, dt \right| < 2\epsilon \qquad \text{for } x \in [x_0, 1).$$

Thus, since ϵ was arbitrary, we conclude that

$$(1 - x) \sum_{n=0}^{\infty} a_n x^n g(x^n) \to c \int_0^1 g(t)\, dt \qquad \text{as } x \to 1.$$

That is, any integrable function that can be "bracketed" in the way described in properties (1)–(3), has the property expressed in (4.6).

It follows from the Weierstrass theorem that any continuous function $f(x)$ on $[0, 1]$ can be "bracketed" in this way, since by that theorem, there exist, for any $\delta > 0$, polynomials $m(x)$ and $M(x)$, which are arbitrarily close to $f(x) - \delta$ and $f(x) + \delta$, respectively, and we have already shown that polynomials satisfy (4.6). That is, continuous functions satisfy (4.6).

Now the special function $h(x)$, whose central role in the proof of Karamata's theorem was discussed earlier, can clearly be "bracketed" between continuous functions in the manner illustrated in Figure 11. That is, $h(x)$ satisfies (4.6), and this completes the proof of Karamata's theorem.

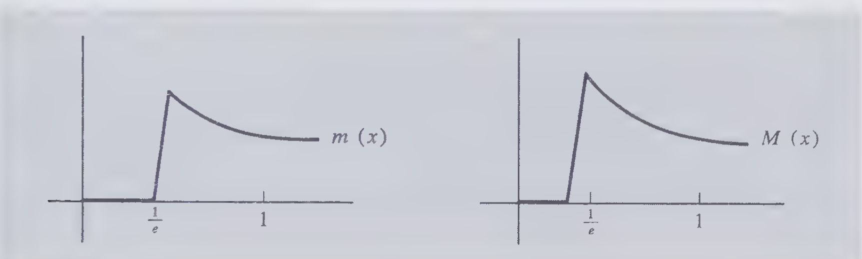

FIGURE 11

Exercises

1 Suppose $f(x)$ is an odd function of class C^1 on $[-\pi, \pi]$ which satisfies $f(-\pi) = f(\pi) = 0$. Show that $f(x)$ can be uniformly approximated by linear combinations of the functions $\sin x, \sin 2x, \sin 3x, \ldots$. (HINT: show that $f'(x)$ must be even, and apply the corresponding result for even functions.)

2 Using the result of problem 1, show that any C^1 function $f(x)$ on $[-\pi, \pi]$ that satisfies $f(-\pi) = f(\pi)$ can be uniformly approximated by linear combinations of the functions $1, \sin x, \cos x, \sin 2x, \cos 2x, \sin 3x, \cos 3x, \ldots$.

3 For a continuous function $f(x)$ on $[0, 1]$, define $F(x) = \int_0^1 e^{tx} f(t)\, dt$. Show that if $F(x) \equiv 0$ on some open interval, then $f(x) \equiv 0$ on $[0, 1]$. (HINT: suppose $F(x) \equiv 0$ on an open interval containing the point x_0. Now $F^{(n)}(x) = \int_0^1 t^n e^{tx} f(t)\, dt$. In particular, setting $x = x_0$, $\int_0^1 t^n e^{tx_0} f(t)\, dt = 0$, for $n = 0, 1, 2, \ldots$.)

CHAPTER FIVE

Fourier Series

1 Orthonormal systems

In this chapter, we discuss an expansion quite different from the power series expansion of a function. We begin with some very general observations.

Suppose $\{f_n(x)\}_{n=1}^{\infty}$ is a sequence of Riemann integrable functions on an interval $[a, b]$. Suppose, moreover that the functions $f_n(x)$ have the following properties:

1. $\displaystyle\int_a^b (f_n(x))^2\,dx = 1, \qquad \text{for } n = 1, 2, \ldots,$

2. $\displaystyle\int_a^b f_k(x)f_n(x)\,dx = 0, \qquad \text{if } k \neq n.$

We call such a sequence *orthonormal.*

Examples

1. Suppose $h(x)$ is defined on the real line by the requirements that for any integer n, positive, negative, or zero,

$$\begin{aligned} h(n) &= h(n + \tfrac{1}{2}) = h(n + 1) = 0; \\ h(x) &= \quad 1, \qquad \text{for } x \in (n, n + \tfrac{1}{2}); \\ h(x) &= -1, \qquad \text{for } x \in (n + \tfrac{1}{2}, n + 1). \end{aligned}$$

Then $h(x)$ is a periodic step-function of period 1. That is, $h(x + 1) = h(x)$.

Define a family of functions $f_1(x), f_2(x), \ldots$ on $[0, 1]$ by setting $f_n(x) = h(2^{n-1}x)$, $x \in [0, 1]$. Then the system $f_1(x), f_2(x), \ldots$ is orthonormal over $[0, 1]$ as the reader can readily verify. This system of functions is known as the system of Rademacher functions. (An understanding of the properties of the Rademacher functions is facilitated by drawing the graphs of the first few functions in the system.)

2. The system

$$\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \sin x, \frac{1}{\sqrt{\pi}} \cos x, \frac{1}{\sqrt{\pi}} \sin 2x, \frac{1}{\sqrt{\pi}} \cos 2x, \ldots$$

is orthonormal over $[-\pi, \pi]$. This follows immediately from the fact that

1. $\int_{-\pi}^{\pi} \sin nx \, dx = \int_{-\pi}^{\pi} \cos nx \, dx = 0, \qquad \text{if } n \geq 1$
2. $\sin^2 nx = \frac{1}{2}(1 - \cos 2nx)$
   $\cos^2 nx = \frac{1}{2}(1 + \cos 2nx)$
3. $\sin mx \cos nx = \frac{1}{2}[\sin (m + n)x + \sin (m - n)x]$
   $\sin mx \sin nx = \frac{1}{2}[\cos (m - n)x - \cos (m + n)x]$
   $\cos mx \cos nx = \frac{1}{2}[\cos (m + n)x + \cos (m - n)x]$

In order to avoid cumbersome formulas involving integrals, it is advantageous at this point to introduce the following notation.

Definition 5.1 For any two functions $f(x)$ and $g(x)$, Riemann integrable over an interval $[a, b]$, define

$$(f, g) = \int_a^b f(x)g(x) \, dx.$$

The following facts are obvious from the definition:

1. $(cf, g) = (f, cg) = c(f, g)$, for any real constant c
2. $(f, g) = (g, f)$
3. $(f + g, h) = (f, h) + (g, h)$
4. $(f, f) \geq 0$

The following definitions are also useful.

Definition 5.2 A function $f(x)$ is said to be a *linear combination* of functions $f_1(x), \ldots, f_n(x)$ if there exist constants $c_1, \ldots, c_n$, such that $f(x) = c_1 f_1(x) + \cdots + c_n f_n(x)$.

Definition 5.3 A collection of functions $f_1(x), \ldots, f_n(x)$ is said to be *linearly independent* if no function in the collection is a linear combination of the remaining functions, or equivalently, if there is no collection of constants $c_1, \ldots, c_n$, not all zero, such that $c_1 f_1(x) + \cdots + c_n f_n(x) \equiv 0$.

We now show how we can construct an orthonormal sequence from a sequence that is not necessarily orthonormal, provided that the latter sequence satisfies certain requirements. More precisely, suppose $\{f_n(x)\}_{n=1}^{\infty}$ is a sequence of continuous functions on an interval $[a, b]$, such that

1. $(f_n, f_n) > 0$, for $n = 1, 2, \ldots,$
2. any finite collection of the functions is linearly independent.

Example

Suppose $[a, b] = [-1, 1]$, and $f_1(x) = 1$, $f_2(x) = x, \ldots, f_n(x) = x^{n-1}, \ldots$. Requirement (1) is obviously satisfied by this system. If requirement (2) were not satisfied, there would exist a polynomial, not all of whose coefficients would be zero, which would vanish identically on $[-1, 1]$. This would contradict the uniqueness of the power series expansion of a function, since the power series for the function that is identically zero on $[-1, 1]$ must be the series consisting of the single constant zero.

Theorem 5.1 *Suppose* $\{f_n(x)\}_{n=1}^{\infty}$ *satisfies requirements* (1) *and* (2) *above. Then there exists a sequence* $\{f_n^*(x)\}_{n=1}^{\infty}$ *of continuous functions on* $[a, b]$, *such that*

A. *The system* $\{f_n^*(x)\}_{n=1}^{\infty}$ *is orthonormal over* $[a, b]$.
B. *For each* n, $f_n^*(x)$ *is a linear combination of* $f_1(x), \ldots, f_n(x)$.
C. *For each* n, $f_n(x)$ *is a linear combination of* $f_1^*(x), \ldots, f_n^*(x)$.

PROOF We shall proceed by induction. Define $b_1 = (f_1, f_1)^{1/2}$; $f_1^*(x) = (1/b_1)f_1(x)$. Then $(f_1^*, f_1^*) = 1$.

Suppose now that orthonormal functions $f_1^*(x), \ldots, f_n^*(x)$ have been found that satisfy requirements (B) and (C). Define a function $g_{n+1}(x)$, by setting

$$g_{n+1}(x) = f_{n+1}(x) - ((f_{n+1}, f_1^*)f_1^*(x) + \cdots + (f_{n+1}, f_n^*)f_n^*(x)).$$

Then $(g_{n+1}, f_k^*) = 0$ for $k = 1, \ldots, n$, since the functions $f_1^*(x), \ldots, f_n^*(x)$ are orthonormal.

Now by (B), the function $(f_{n+1}, f_1^*)f_1^*(x) + \cdots + (f_{n+1}, f_n^*)f_n^*(x)$ can be expressed as a linear combination of the functions $f_1(x), \ldots, f_n(x)$, and this implies that $g_{n+1}(x)$ can be expressed as a linear combination of the functions $f_1(x), \ldots, f_{n+1}(x)$. Since the f_k are linearly independent, this implies that $g_{n+1}(x) \not\equiv 0$, and hence $(g_{n+1}, g_{n+1}) \neq 0$, since $g_{n+1}(x)$ is clearly continuous. Define $b_{n+1} = (g_{n+1}, g_{n+1})^{1/2}$, and set

$$f_{n+1}^*(x) = \frac{1}{b_{n+1}} g_{n+1}(x).$$

Then the functions $f_1^*(x), \ldots, f_{n+1}^*(x)$ are orthonormal and clearly satisfy (B) and (C), with n replaced by $n + 1$, and this completes the induction.

The sequence $\{f_n^*(x)\}_{n=1}^{\infty}$ obtained in the above way is often called the sequence obtained by *orthonormalizing* the sequence $\{f_n(x)\}_{n=1}^{\infty}$.

Example

The polynomials obtained by orthonormalizing $1, x, x^2, \ldots$ over $[-1, 1]$ are known as the normalized Legendre polynomials.

Suppose now that $\{f_n(x)\}_{n=1}^{\infty}$ is an orthonormal sequence on an interval $[a, b]$, and suppose $g(x)$ is Riemann integrable over $[a, b]$. We wish to investigate the possibility of expanding $g(x)$ in a series of the form

$$\sum_{n=1}^{\infty} a_n f_n(x). \tag{5.1}$$

How should we choose the a_n's? At this point, the requirement that $\{f_n(x)\}_{n=1}^{\infty}$ is orthonormal suggests an answer. For suppose such an expansion existed. Then, at least on a purely formal basis, without regard to the validity of the operations involved, we should have, for any integer $k \geq 1$,

$$\begin{aligned}(g, f_k) &= \sum_{n=1}^{\infty} a_n(f_n, f_k)\\ &= a_k.\end{aligned}$$

For the moment, we shall not worry about whether or not the expansion (5.1) with $a_n = (g, f_n)$ converges to $g(x)$. Such questions can be extremely delicate. As this preliminary stage, we merely wish to show that if we define the a_n's in this way, then a certain inequality must always be true.

Theorem 5.2 **(Bessel's Inequality)** *Suppose $\{f_n(x)\}_{n=1}^{\infty}$ is an orthonormal sequence over $[a, b]$, and suppose $g(x)$ is Riemann integrable over $[a, b]$. Define $a_n = (g, f_n)$. Then the series $\sum_{n=1}^{\infty} a_n^2$ is convergent. Moreover $\sum_{n=1}^{\infty} a_n^2 \leq (g, g)$.*

PROOF It clearly suffices to show that for each positive integer N, $\sum_{n=1}^{N} a_n^2 \leq (g, g)$. In order to show this, consider the quantity

$$\left(g - \sum_{n=1}^{N} a_n f_n, g - \sum_{n=1}^{N} a_n f_n\right),$$

which is ≥ 0. Now

$$\left(g - \sum_{n=1}^{N} a_n f_n, g - \sum_{n=1}^{N} a_n f_n\right)$$
$$= (g, g) - 2\sum_{n=1}^{N} (f_n, g)a_n + \left(\sum_{n=1}^{N} a_n f_n, \sum_{n=1}^{N} a_n f_n\right),$$

But $2\sum_{n=1}^{N} (f_n, g)a_n = 2\sum_{n=1}^{N} a_n^2$, and $(\sum_{n=1}^{N} a_n f_n, \sum_{n=1}^{N} a_n f_n) = \sum_{n=1}^{\infty} a_n^2$, since the sequence $\{f_n(x)\}_{n=1}^{\infty}$ is orthonormal.

We conclude therefore that

$$\left(g - \sum_{n=1}^{N} a_n f_n, g - \sum_{n=1}^{N} a_n f_n\right) = (g, g) - 2\sum_{n=1}^{N} a_n^2 + \sum_{n=1}^{N} a_n^2$$
$$= (g, g) - \sum_{n=1}^{N} a_n^2.$$

Now $(g - \sum_{n=1}^{N} a_n f_n, \; g - \sum_{n=1}^{N} a_n f_n) \geq 0$, so $(g, g) - \sum_{n=1}^{N} a_n^2 \geq 0$, or $(g, g) \geq \sum_{n=1}^{N} a_n^2$, which completes the proof.

Theorem 5.3 (The Theorem of Best Approximation in the Mean) *Suppose $\{f_n(x)\}_{n=1}^{\infty}$ is an orthonormal sequence on $[a, b]$, $g(x)$ is a Riemann integrable function on $[a, b]$, and $N > 0$ is an integer. Then the quantity*

$$\left(g - \sum_{n=1}^{N} a_n f_n, g - \sum_{n=1}^{N} a_n f_n\right) \tag{5.2}$$

is minimized when $a_j = (g, f_j)$, $j = 1, \ldots, N$. In fact, any alternative choice of the constants $a_1, \ldots, a_N$ makes the quantity (5.2) strictly larger.

PROOF For any $a_1, \ldots, a_N$,

$$\left(g - \sum_{n=1}^{N} a_n f_n, g - \sum_{n=1}^{N} a_n f_n\right) = (g, g) - 2\sum_{n=1}^{N} a_n(g, f_n) + \sum_{n=1}^{N} a_n^2$$
$$= (g, g) - \sum_{n=1}^{N} (g, f_n)^2 + \sum_{n=1}^{N} (a_n - (g, f_n)^2)^2,$$

and this implies the desired result.

Theorem 5.4 (Cauchy-Schwarz) *Suppose $f(x)$ and $g(x)$ are Riemann integrable over an interval $[a, b]$. Then*

$$|(f, g)| \leq (f, f)^{1/2}(g, g)^{1/2}.$$

PROOF The assertion is, of course, equivalent to the statement that $(f, g)^2 \leq (f, f)(g, g)$, and it is in this form that we shall prove it.

Supposing $f(x)$ and $g(x)$ fixed, consider, as a function of t, the quantity $(f - tg, f - tg)$, where $t \in (-\infty, \infty)$. Now $(f - tg, f - tg) \geq 0$, and if we expand the quantity on the left, we find that $(f, f) - 2t(f, g) + t^2(g, g) \geq 0$. Now the graph of $(f, f) - 2t(f, g) + t^2(g, g)$, regarded as a function of t, is either a straight line or a parabola, depending on whether or not (g, g) vanishes. In the first case, it is clear that (f, g) must be zero, since the straight line can never pass below the t axis, and if $(f, g) = 0$, the theorem is obvious. Accordingly, we assume that $(g, g) > 0$. Now since

$$(f, f) - 2t(f, g) + t^2(g, g) \geq 0$$

for all t, we conclude that there exists at most one value of t for which $(f, f) - 2t(f, g) + t^2(g, g)$ is zero.

Now the zeros of a polynomial of the form $at^2 + bt + c$ are given by

$$\frac{-b + (b^2 - 4ac)^{1/2}}{2a} \quad \text{and} \quad \frac{-b - (b^2 - 4ac)^{1/2}}{2a},$$

respectively, and from this it is clear that if $b^2 - 4ac$ is > 0, then the polynomial $at^2 + bt + c$ has two distinct real zeros. Now as we have seen, the polynomial $(f, f) - 2t(f, g) + t^2(g, g)$ has at most one real zero, so we conclude that $4(f, g)^2 - 4(f, f)(g, g) \leq 0$, or $(f, g)^2 \leq (f, f)(g, g)$, which proves the theorem.

One very important consequence of the last theorem is the following.

Theorem 5.5 (Minkowski's Inequality) *Suppose $f(x)$ and $g(x)$ are functions defined over $[a, b]$. Then*

$$\left[\int_a^b (f(x) + g(x))^2\, dx\right]^{1/2} \leq \left[\int_a^b (f(x))^2\, dx\right]^{1/2} + \left[\int_a^b (g(x))^2\, dx\right]^{1/2}$$

That is,

$$(f + g, f + g)^{1/2} \leq (f, f)^{1/2} + (g, g)^{1/2}.$$

PROOF By squaring both sides, we see that the assertion is equivalent to the statement that

$$(f + g, f + g) \leq (f, f) + 2(f, f)^{1/2}(g, g)^{1/2} + (g, g),$$

or

$$(f, f) + 2(f, g) + (g, g) \leq (f, f) + 2(f, f)^{1/2}(g, g)^{1/2} + (g, g),$$

or

$$(f, g) \leq (f, f)^{1/2}(g, g)^{1/2}.$$

But in the last form, the assertion is implied by the Cauchy-Schwarz inequality.

2 The notion of completeness

Definition 5.4 An orthonormal system $\{f_n(x)\}$ over an interval $[a, b]$ is said to be *complete* over $[a, b]$ if, for any continuous function $g(x)$ on $[a, b]$, there exists a sequence $\{L_n(x)\}$ of linear combination of the f_n's, such that $(g - L_n, g - L_n) \to 0$ as $n \to \infty$.

Example

The normalized Legendre polynomials are complete over $[-1, 1]$. More generally, if $f(x)$ is a continuous function and is strictly increasing or strictly decreasing over an interval $[a, b]$, then the system obtained by orthonormalizing $1, f(x), (f(x))^2, (f(x))^3, \ldots$ is complete over $[a, b]$.

To see this, note that by the generalization of the Weierstrass approximation theorem, there exists, for any continuous function $g(x)$, a sequence $\{L_n(x)\}$ of linear combinations of the functions $1, f(x), (f(x))^2, (f(x))^3, \ldots$, such that $L_n(x) \to g(x)$ uniformly over $[a, b]$. From this, it is clear that $(g - L_n, g - L_n) \to 0$. Now any linear combination of the functions $1, f(x), (f(x))^2, (f(x))^3, \ldots$ can be expressed as a linear combination of the functions in the sequence obtained by orthonormalizing $1, f(x), (f(x))^2, (f(x))^3, \ldots$, so the result follows immediately.

Theorem 5.6 *If $\{f_n(x)\}_{n=1}^{\infty}$ is a* ***complete*** *orthonormal system over an interval $[a, b]$, then the only continuous function $g(x)$ for which $(g, f_n) = 0$ for all $n \geq 1$ is the function that is identically zero on $[a, b]$.*

PROOF By hypothesis, there exists a sequence $\{L_n(x)\}$ of linear combinations of the f_n's, such that $(g - L_n, g - L_n) \to 0$, or $(g, g) + (L_n, L_n) - 2(g, L_n) \to 0$. Now if $(g, f_n) = 0$ for all n, this implies that $(g, g) + (L_n, L_n) \to 0$, which clearly implies that $(g, g) = 0$, since both (g, g) and (L_n, L_n) are ≥ 0. This in turn clearly implies that $g(x) = 0$ over $[a, b]$, which completes the proof.

It is an extremely important fact that Bessel's inequality becomes an equality for complete orthonormal systems.

Theorem 5.7 (Parseval's Equality) *Suppose $\{f_n(x)\}$ is a complete orthonormal system over $[a, b]$, and suppose $g(x)$ is Riemann integrable on $[a, b]$. Define $a_n = (g, f_n)$, $n = 1, 2, 3, \ldots$. Then $\sum_{n=1}^{\infty} a_n^2 = (g, g)$.*

PROOF We already know from Bessel's inequality that the series $\sum_{n=1}^{\infty} a_n^2$ converges, and that the limit is $\leq (g, g)$. It remains to show that if $\{f_n(x)\}$ is complete, then the limit is, in fact, (g, g). Now since $\{f_n(x)\}$ is complete, there exist, for every $\epsilon > 0$, constants $b_1, \ldots, b_N$, such that

$$\left(g - \sum_{n=1}^{N} b_n f_n, g - \sum_{n=1}^{N} b_n f_n\right) \leq \epsilon.$$

By the theorem of best approximation in the mean, the expression

$$\left(g - \sum_{n=1}^{N} b_n f_n, g - \sum_{n=1}^{N} b_n f_n\right)$$

is minimized if $b_n = (g, f_n)$. That is, if $b_n = a_n$. From this we conclude that

$$\left(g - \sum_{n=1}^{N} a_n f_n, g - \sum_{n=1}^{N} a_n f_n\right) \leq \epsilon.$$

But as we remarked in the proof of Bessel's inequality,

$$\left(g - \sum_{n=1}^{N} a_n f_n, g - \sum_{n=1}^{N} a_n f_n\right) = (g, g) - \sum_{n=1}^{N} a_n^2.$$

Since by Bessel's inequality the last quantity is ≥ 0, it follows that

$$0 \leq (g, g) - \textstyle\sum_{n=1}^{N} a_n^2 \leq \epsilon.$$

From this it is clear that $0 \leq (g, g) - \sum_{n=1}^{\infty} a_n^2 \leq \epsilon$ (the left side of the last inequality is simply Bessel's inequality), and since ϵ was arbitrary, it follows that $\sum_{n=1}^{\infty} a_n^2 = (g, g)$.

3 Some lemmas

In this section, we derive some results that pave the way for a detailed discussion of the trigonometric system in the next section.

Lemma 5.1 *Suppose $f(x)$ is defined for $-\infty < x < \infty$ and is periodic of period α. That is, $f(x + \alpha) = f(x)$, for any x. Suppose, moreover, that $f(x)$ is Riemann integrable over any finite interval. Then the function $F(x)$, defined by $F(x) = \int_x^{x+\alpha} f(t)\, dt$, is constant.*

PROOF It suffices to show that for any two numbers x_1 and x_2,

$$\int_{x_1}^{x_1+\alpha} f(t)\,dt - \int_{x_2}^{x_2+\alpha} f(t)\,dt = 0.$$

Now

$$\begin{aligned}\int_{x_1}^{x_1+\alpha} f(t)\,dt - \int_{x_2}^{x_2+\alpha} f(t)\,dt &= \int_{x_1}^{x_1+\alpha} f(t)\,dt - \left(\int_{x_1+\alpha}^{x_2+\alpha} f(t)\,dt - \int_{x_1+\alpha}^{x_2} f(t)\,dt\right)\\ &= \int_{x_1}^{x_1+\alpha} f(t)\,dt + \int_{x_1+\alpha}^{x_2} f(t)\,dt - \int_{x_1+\alpha}^{x_2+\alpha} f(t)\,dt\\ &= \int_{x_1}^{x_2} f(t)\,dt - \int_{x_1+\alpha}^{x_2+\alpha} f(t)\,dt.\end{aligned}$$

But the substitution $t^* = t - \alpha$ shows that the second integral equals the first, if we bear in mind that $f(t + \alpha) = f(t)$, and this proves the lemma.

Lemma 5.2 *Suppose $f(x)$ is C^2 on an interval $[-\epsilon, \epsilon]$, where $\epsilon < 2\pi$, and suppose $f(0) = 0$ and $f'(0) = a$. Then the function $h(x)$, defined by*

$$h(x) = \frac{f(x)}{\sin x/2}, \qquad \text{for } x \in [-\epsilon, \epsilon]; \quad x \neq 0$$

$$h(0) = 2a,$$

is C^1 on $[-\epsilon, \epsilon]$.

PROOF By Taylor's theorem,

$$\sin\frac{x}{2} = \frac{x}{2}\left[1 - \frac{1}{3!}\left(\frac{x}{2}\right)^2 + \frac{1}{5!}\left(\frac{x}{2}\right)^4 - \cdots\right],$$

and since $\sin x/2$ only vanishes at the point $x = 0$ of the interval $[-\epsilon, \epsilon]$, we conclude that the function

$$H(x) = 1 - \frac{1}{3!}\left(\frac{x}{2}\right)^2 + \frac{1}{5!}\left(\frac{x}{2}\right)^4 - \cdots,$$

is never zero on $[-\epsilon, \epsilon]$, has the value 1 at $x = 0$, and is, of course, C^1 on $[-\epsilon, \epsilon]$.

Now $f(x) = \int_0^x f'(t)\,dt$ for $x \in [-\epsilon, \epsilon]$, since $f(0) = 0$. But

$$\int_0^x f'(t)\,dt = x\int_0^1 f'(xt)\,dt, \qquad \text{so} \quad \frac{f(x)}{\sin x/2} = \frac{2}{H(x)}\int_0^1 f'(xt)\,dt.$$

Now it is easily verified (for example, by integration) that the function $g(x) = \int_0^1 f'(xt)\,dt$ is C^1 on $[-\epsilon, \epsilon]$, $g'(x) = \int_0^1 tf''(xt)\,dt$, and $g(0) = a$. The lemma follows from this.

Lemma 5.3 (Riemann-Lebesgue) *Suppose $f(x)$ is piecewise C^1 on an interval $[a, b]$. That is, $[a, b]$ can be subdivided into a finite number of intervals $[x_0, x_1], \ldots, [x_{n-1}, x_n]$, with $x_0 = a$ and $x_n = b$, and such that on each of these subintervals $f(x)$ is C^1, after modification at the endpoints, if necessary. Then the integral $\int_a^b f(x) \sin xy\, dx$ tends to zero as $y \to \infty$. In fact, there exists an $M > 0$, such that*

$$\left|\int_a^b f(x) \sin xy\, dx\right| \leq \frac{M}{|y|}.$$

PROOF

$$\int_a^b f(x) \sin xy\, dx = \int_{x_0}^{x_1} f(x) \sin xy\, dx + \cdots + \int_{x_{n-1}}^{x_n} f(x) \sin xy\, dx,$$

so it clearly suffices to prove the result for a typical integral on the right.

Now for $y \neq 0$,

$$\int_{x_j}^{x_{j+1}} f(x) \sin xy\, dx = -\frac{1}{y} f(x) \cos xy \Big|_{x=x_j}^{x=x_{j+1}} + \frac{1}{y}\int_{x_j}^{x_{j+1}} f'(x) \cos xy\, dx$$

if we integrate by parts, and it is clear that the last two quantities tend to zero like $1/y$ as $y \to \infty$.

Definition 5.5 The *Dirichlet kernel* of order n, denoted by $D_n(x)$, is defined to be the function $\frac{1}{2} + \cos x + \cdots + \cos nx$ if $n > 0$, and $\frac{1}{2}$ if $n = 0$. Note that $D_n(x)$ is analytic on $(-\infty, \infty)$ and periodic of period 2π, since it is a finite sum of functions having this property.

Lemma 5.4 $\displaystyle \frac{1}{\pi}\int_{-\pi}^{\pi} D_n(x)\, dx = 1, \quad n = 0, 1, 2, \ldots.$

PROOF This follows immediately from the fact that

$$\int_{-\pi}^{\pi} \cos kx\, dx = 0, \qquad \text{for any integer } k \geq 1.$$

Lemma 5.5 $D_n(x) = [\sin (n + \frac{1}{2})x]/[2 \sin x/2]$, *if we adopt the convention that the value of the function on the right is $n + \frac{1}{2}$ for $x = 0, \pm 2\pi, \pm 4\pi, \ldots$. That is, for those points at which the denominator vanishes.*

PROOF The assertion is obvious for $x = 0, \pm 2\pi, \pm 4\pi, \ldots$, so we may assume $\sin x/2 \neq 0$. Now the assertion is equivalent to the statement that

$$\cos x + \cdots + \cos nx = \frac{\sin (n + \frac{1}{2})x - \sin \frac{1}{2}x}{2 \sin \frac{1}{2}x}$$

or that

$$2 \cos x \sin \tfrac{1}{2}x + \cdots + 2 \cos nx \sin \tfrac{1}{2}x = \sin (n + \tfrac{1}{2})x - \sin \tfrac{1}{2}x.$$

Now by the trigonometric addition formulas,

$$2 \cos kx \sin \tfrac{1}{2}x = \sin (k + \tfrac{1}{2})x - \sin (k - \tfrac{1}{2})x.$$

That is,

$$\begin{aligned} 2 \cos x \sin \tfrac{1}{2}x + \cdots + 2 \cos nx \sin \tfrac{1}{2}x &= (\sin \tfrac{3}{2}x - \sin \tfrac{1}{2}x) + (\sin \tfrac{5}{2} - \sin \tfrac{3}{2}x) + \cdots \\ &\quad + (\sin (n + \tfrac{1}{2})x - \sin (n - \tfrac{1}{2})x) \\ &= \sin (n + \tfrac{1}{2})x - \sin \tfrac{1}{2}x, \end{aligned}$$

since all other terms cancel out, and this proves the Lemma.

4 The trigonometric system

In this section, we investigate one of the most important orthonormal systems, namely the system

$$\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \sin x, \frac{1}{\sqrt{\pi}} \cos x, \frac{1}{\sqrt{\pi}} \sin 2x, \frac{1}{\sqrt{\pi}} \cos 2x, \ldots.$$

This system, as we have already observed in Section 1, is orthonormal over $[-\pi, \pi]$ and, in fact, over any interval of the form $[a, a + 2\pi]$.

Suppose $g(x)$ is Riemann integrable over $[-\pi, \pi]$. Following Section 1 we are led to consider the series

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx), \tag{5.3}$$

where

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} g(x)\, dx,$$

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} g(x) \cos nx\, dx, \qquad \text{and} \qquad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} g(x) \sin nx\, dx.$$

The series (5.3) is called the *Fourier series* of the function $g(x)$. (To simplify the notation, we have omitted the factor $1/\sqrt{\pi}$ from the functions in the orthonormal system, but this is compensated for by an additional factor of $1/\sqrt{\pi}$, which we have included in the definition of the a_n's and b_n's.

Remark In the definition of a_0, we could avoid the rather artificial-looking factor $\frac{1}{2}$ by initially defining a_0 to be $(1/2\pi)\int_{-\pi}^{\pi} g(x)\,dx$, rather than $(1/\pi)\int_{-\pi}^{\pi} g(x)\,dx$. The only good reason for defining a_0 as we have, is that the integral formulas that give the a_i's and b_i's are all preceded by the same factor of $1/\pi$, which presumably makes them easier to remember as a whole. This is, of course, strictly a matter of taste, although it does have the advantage of being the usual definition.

Now for a given x, the convergence of (5.3) is equivalent to the convergence of the associated sequence $\{s_n(x)\}_{n=0}^{\infty}$ of partial sums, where

$$s_0(x) = \frac{a_0}{2};$$

$$s_N(x) = \frac{a_0}{2} + \sum_{n=1}^{N} (a_n \cos nx + b_n \sin nx), \qquad \text{for } N \geq 1.$$

Let us consider these partial sums more closely. By definition, for $N \geq 1$,

$$\begin{aligned}
s_N(x) &= \frac{a_0}{2} + \sum_{n=1}^{N} (a_n \cos nx + b_n \sin nx) \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} \tfrac{1}{2} g(t)\,dt \\
&\qquad + \sum_{n=1}^{N} \left(\frac{1}{\pi}\int_{-\pi}^{\pi} \cos nx g(t) \cos nt\,dt + \frac{1}{\pi}\int_{-\pi}^{\pi} \sin nx g(t) \sin nt\,dt \right) \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} [\tfrac{1}{2} + (\cos x \cos t + \sin x \sin t) + \cdots \\
&\qquad\qquad + (\cos nx \cos nt + \sin nx \sin nt)] g(t)\,dt \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} [\tfrac{1}{2} + \cos (t - x) + \cdots + \cos n(t - x)] g(t)\,dt \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} D_n(t - x) g(t)\,dt,
\end{aligned}$$

where D_n is the Dirichlet kernel.

Since the $s_n(x)$ are periodic of period 2π, it is evident that if they converge at a point x, they must also converge, and to the same value, at the points $x \pm 2\pi$, $x \pm 4\pi$, $x \pm 6\pi, \ldots$. For this reason, it is natural to suppose that our function $g(x)$ is, in fact, defined on $(-\infty, \infty)$ and is periodic of period 2π. If $g(x)$ is initially only defined for $x \in [-\pi, \pi]$, with $g(-\pi) = g(\pi)$, we can

obviously extend it in exactly one way to $(-\infty, \infty)$, if we impose the requirement that the result be periodic of period 2π.

We can now state and prove the main result of this chapter.

Theorem 5.8 *Suppose $g(x)$ is periodic of period 2π over $(-\infty, \infty)$. Moreover, suppose that $g(x)$ is piecewise C^1 on $[-\pi, \pi]$ and that $g(x)$ is C^2 in a neighborhood of the point x_0. Then the Fourier series for $g(x)$ converges to $g(x_0)$ at the points $x = x_0$.*

PROOF We must show that $s_N(x_0) - g(x_0) \to 0$ *as* $N \to \infty$. Now

$$
\begin{aligned}
s_N(x_0) &= \frac{1}{\pi}\int_{-\pi}^{\pi} D_n(t - x_0)g(t)\,dt \\
&= \frac{1}{\pi}\int_{-\pi-x_0}^{\pi-x_0} D_n(t)g(t + x_0)\,dt \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} D_n(t)g(t + x_0)\,dt
\end{aligned}
$$

by Lemma 5.1. On the other hand, $g(x_0) = (1/\pi)\int_{-\pi}^{\pi} D_n(t)g(x_0)\,dt$, since $(1/\pi)\int_{-\pi}^{\pi} D_n(t)\,dt = 1$, for $n = 0, 1, 2, \ldots$. Thus,

$$
s_N(x_0) - g(x_0) = \frac{1}{\pi}\int_{-\pi}^{\pi} D_n(t)(g(t + x_0) - g(x_0))\,dt.
$$

Now by hypothesis, $g(t + x_0) - g(x_0)$ is C^2 in a small interval $[-\epsilon, \epsilon]$ and is equal to zero for $t = 0$. Thus, by Lemma 5.2,

$$
\frac{g(t + x_0) - g(x_0)}{2 \sin \frac{1}{2}t}
$$

has a C^1 extension to $[-\epsilon, \epsilon]$, which we shall call $h(t)$. It is, moreover, clear that if we extend $h(t)$ to $[-\pi, \pi]$ by defining

$$
h(t) = \frac{g(t + x_0) - g(x_0)}{2 \sin \frac{1}{2}t},
$$

the result is piecewise C^1 over $[-\pi, \pi]$. Thus,

$$
\begin{aligned}
s_N(x_0) - g(x_0) &= \frac{1}{\pi}\int_{-\pi}^{\pi} D_n(t)(g(t + x_0) - g(x_0))\,dt \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} h(t) \sin (n + \tfrac{1}{2})t\,dt,
\end{aligned}
$$

and by Lemma 5.3, the last quantity tends to zero as $n \to \infty$, which proves the theorem.

Remark There are stronger versions of Theorem 5.8. For example, it is known that if a function exhibits one of certain simple types of discontinuities at a point, then the Fourier series for the function converges to the average of the left and right limits of the function at that point (*cf.* [1], p. 594, and problem 7 at the end of this chapter).

Corollary of Theorem 5.8 *Suppose $g(x)$ is C^2 on $(-\infty, \infty)$ and periodic of period 2π. Then the Fourier series for $g(x)$ converges uniformly to $g(x)$ over $[-\pi, \pi]$, and hence over $(-\infty, \infty)$, since all functions involved are periodic of period 2π.*

PROOF The last theorem shows that the Fourier series for $g(x)$ converges to $g(x)$ at all points. It remains to show that the convergence is uniform. To see this, recall that for $n \geq 1$,

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} g(x)\cos nx\,dx, \qquad \text{and} \qquad b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} g(x)\sin nx\,dx.$$

Now if we integrate by parts twice, using the fact that $g(x)$, and hence $g'(x)$, is periodic of period 2π, we find that there exists an $M > 0$, such that $|a_n|, |b_n| \leq M/n^2$. This implies that the partial sums of

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

are dominated by the terms of an absolutely convergent series of constants, which clearly implies the desired result.

By using a somewhat more involved argument, we can prove the following generalization of the last theorem.

Theorem 5.9 *Suppose $g(x)$ is continuous on $(-\infty, \infty)$ and periodic of period 2π. Suppose, moreover, that on $[-\pi, \pi]$, $g(x)$ is piecewise C^2. That is, $[-\pi, \pi]$ can be subdivided into a finite number of intervals $[x_0, x_1], \ldots, [x_{n-1}, x_n]$, with $x_0 = -\pi$ and $x_n = \pi$, such that on each of these subintervals $g(x)$ is C^2. Then the Fourier series for $g(x)$ converges uniformly to $g(x)$ on $[-\pi, \pi]$, and hence on $(-\infty, \infty)$, since all functions involved are periodic of period 2π.*

Remark This implies, among other things, that if $g(x)$ is a polygonal function defined on $[-\pi, \pi]$, and such that $g(-\pi) = g(\pi)$, then the Fourier series for $g(x)$ converges uniformly to $g(x)$.

PROOF Denote the Fourier coefficients of $g(x)$ by $a_0, a_1, a_2, \ldots; b_1, b_2, \ldots$, and those of $g'(x)$ by $a'_0, a'_1, a'_2, \ldots; b'_1, b'_2, \ldots$. To prove the theorem, it certainly suffices to show that both $\sum_{n=0}^{\infty} |a_n|$ and $\sum_{n=1}^{\infty} |b_n|$ are convergent. We shall prove this for the first series. The proof for the second is essentially the same.

Now

$$\begin{aligned} a_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} g(x)\cos kx\,dx \\ &= \frac{1}{\pi}\int_{x_0}^{x_1} g(x)\cos kx\,dx + \cdots + \frac{1}{\pi}\int_{x_{n-1}}^{x_n} g(x)\cos kx\,dx. \end{aligned}$$

But integrating by parts,

$$\frac{1}{\pi}\int_{x_j}^{x_{j+1}} g(x)\cos kx\,dx = \frac{1}{\pi k} g(x)\sin kx\Big|_{x_j}^{x_{j+1}} - \frac{1}{\pi k}\int_{x_j}^{x_{j+1}} g'(x)\sin kx\,dx.$$

Now

$$\frac{1}{\pi k} g(x)\sin kx\Big|_{x_0}^{x_1} + \cdots + \frac{1}{\pi k} g(x)\sin kx\Big|_{x_{n-1}}^{x_n} = 0,$$

since $g(x_0)\sin kx_0 = g(x_n)\sin kx_n$, and the remaining summands obviously cancel out in pairs, since $g(x)$ is continuous. Thus,

$$\begin{aligned} &\frac{1}{\pi}\int_{x_0}^{x_1} g(x)\cos kx\,dx + \cdots + \frac{1}{\pi}\int_{x_{n-1}}^{x_n} g(x)\cos kx\,dx \\ &\quad= -\frac{1}{\pi k}\left(\int_{x_0}^{x_1} g'(x)\sin kx\,dx + \cdots + \int_{x_{n-1}}^{x_n} g'(x)\sin kx\,dx\right) \\ &\quad= -\frac{1}{\pi k}\int_{x_0}^{x_1} g'(x)\sin kx\,dx \\ &\quad= -\frac{1}{k} a'_k \end{aligned}$$

or

$$a_k = -\frac{1}{k} a'_k.$$

Now by Bessel's inequality, $\sum_{k=1}^{\infty} (a'_k)^2$ is convergent.

We wish to show that $\sum_{k=1}^{\infty} |a_k|$ is convergent. To do this, we simply note that $\sum_{k=1}^{\infty} |a_k|$ can be expressed as $\sum_1 |a_k| + \sum_2 |a_k|$, where the first sum is taken over all those a_k's for which $|a_k| \leq |a'_k|^2$, and the second sum is taken over all those a_k's for which $|a_k| > |a'_k|^2$, that is, for which $(1/k)\,|a'_k| > |a'_k|^2$ or $1/k > |a'_k|$.

The first sum clearly converges by the majorization test. To see that the second sum converges, note that $|a_k| = (1/k)|a_k'|$, so all a_k's in the second sum satisfy the inequality $|a_k| < 1/k^2$, and hence the second sum also converges by the majorization test, which completes the proof.

The last theorem leads to an alternative proof of the Weierstrass approximation theorem. Suppose $f(x)$ is continuous on an interval $[a, b]$. Given $\epsilon > 0$, we wish to show that there exists a polynomial $P(x)$, such that $|f(x) - P(x)| \leq \epsilon$ for $x \in [a, b]$. It suffices to establish this for the case in which $a = -1$ and $b = 1$, since the result for a general interval follows from the result for $[-1, 1]$ by a transformation of the form $x \to \alpha x + \beta$. Accordingly, suppose $f(x)$ is continuous on $[-1, 1]$. Then there certainly exists a polygonal function $g(x)$ on $[-\pi, \pi]$, such that $|f(x) - g(x)| \leq \frac{1}{3}\epsilon$ for $x \in [-1, 1]$, and $g(-\pi) = g(\pi) = 0$ (Figure 12).

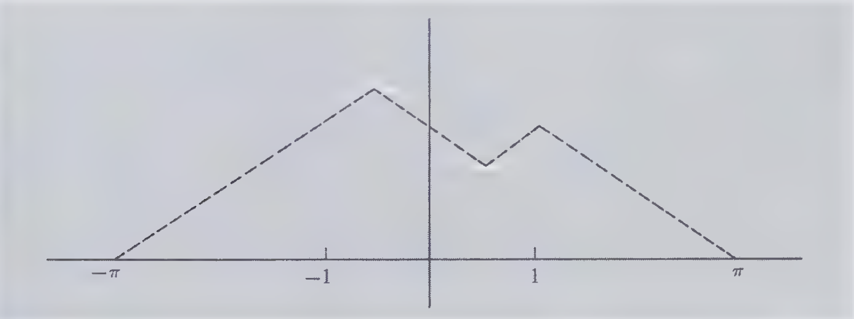

FIGURE 12

By the last theorem, there exist numbers $a_0, \ldots, a_N$ and $b_1, \ldots, b_N$, such that

$$\left| g(x) - \frac{a_0}{2} - \sum_{n=1}^{N} (a_n \cos nx + b_n \sin nx) \right| \leq \tfrac{1}{3}\epsilon$$

for $x \in [-\pi, \pi]$. Thus, by the triangle inequality,

$$\left| f(x) - \frac{a_0}{2} - \sum_{n=1}^{N} (a_n \cos nx + b_n \sin nx) \right| \leq \tfrac{2}{3}\epsilon.$$

Now the power series for $a_0/2 + \sum_{n=1}^{N} (a_n \cos nx + b_n \sin nx)$ converges uniformly on any finite interval, so there clearly exists a polynomial $P(x)$, such that

$$\left| \frac{a_0}{2} + \sum_{n=1}^{N} (a_n \cos nx + b_n \sin nx) - P(x) \right| \leq \tfrac{1}{3}\epsilon$$

for $x \in [-1, 1]$. Again invoking the triangle inequality we find that

$$|f(x) - P(x)| \le \epsilon \qquad \text{for } x \in [-1, 1],$$

which was to be proved.

Definition 5.6 By a *trigonometric polynomial*, we mean a function of the form

$$\frac{a_0}{2} + \sum_{n=1}^{N} (a_n \cos nx + b_n \sin nx),$$

for some constants $a_0, \ldots, a_N$, and $b_1, \ldots, b_N$.

The following result is an extremely important consequence of Theorem 5.9.

Theorem 5.10 *The trigonometric system is complete over* $[-\pi, \pi]$.

PROOF Recall that completeness is equivalent to the statement that for any $\epsilon > 0$ and any continuous function $g(x)$ on $[-\pi, \pi]$, there exists a trigonometric polynomial $T(x)$, such that

$$\left[\int_{-\pi}^{\pi} (g(x) - T(x))^2\, dx\right]^{1/2} \le \epsilon.$$

In order to show that this is the case, we simply note that there certainly exists a polygonal function $h(x)$ on $[-\pi, \pi]$, such that $h(-\pi) = h(\pi)$, and

$$\left[\int_{-\pi}^{\pi} (g(x) - h(x))^2\, dx\right]^{1/2} \le \tfrac{1}{2}\epsilon.$$

We could, for example, define $h(x)$ in such a way that $|g(x) - h(x)|$ is very small over an interval of the form $[-(\pi - \delta), (\pi - \delta)]$, while on the interval $[-\pi, -(\pi - \delta)]$, the graph of $h(x)$ is a straight line segment connecting the point $(-\pi, 0)$ to the point $(-(\pi - \delta), g(-(\pi - \delta)))$, and on the interval $[\pi - \delta, \pi]$, the graph is a straight line segment connecting the points $(\pi - \delta, g(\pi - \delta))$ and $(\pi, 0)$ (see Figure 13).

By taking δ small enough, and at the same time making certain that $|g(x) - h(x)|$ is sufficiently small over $[-(\pi - \delta), \pi - \delta]$, we can make the integral $[\int_{-\pi}^{\pi} (g(x) - h(x))^2\, dx]^{1/2}$ as small as we please.

Accordingly, suppose that we have found a polygonal function $h(x)$, such that

$$\left[\int_{-\pi}^{\pi} (g(x) - h(x))^2\, dx\right]^{1/2} \le \tfrac{1}{2}\epsilon.$$

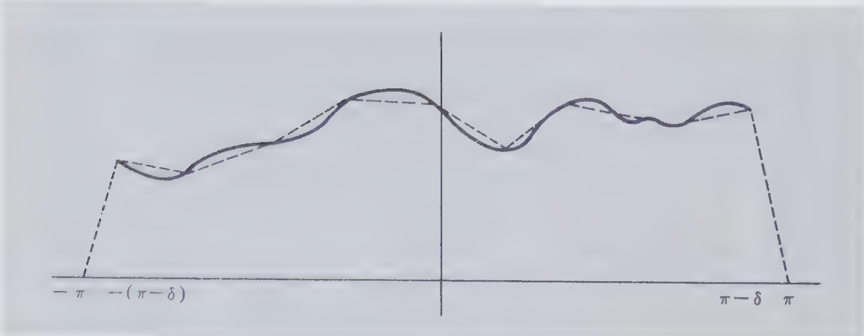

FIGURE 13

By Theorem 5.9, there exists a trigonometric polynomial $T(x)$, such that

$$\left[\int_{-\pi}^{\pi} (h(x) - T(x))^2\, dx\right]^{1/2} \le \tfrac{1}{2}\epsilon,$$

so it follows from Minkowski's inequality that

$$\left[\int_{-\pi}^{\pi} (g(x) - T(x))^2\, dx\right]^{1/2} \le \epsilon,$$

which proves the theorem.

Corollary *Parseval's equality holds for the trigonometric system. That is, if $g(x)$ is a continuous function over $[-\pi, \pi]$, with Fourier series*

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx),$$

then

$$\frac{1}{\pi}\int_{-\pi}^{\pi} (g(x))^2\, dx = \left(\frac{a_0}{2}\right)^2 + \sum_{n=1}^{\infty} (a_n^2 + b_n^2).$$

Remark Note that there is no requirement that $g(x)$ be periodic. That is, for Parseval's equality, it is not necessary that $g(-\pi) = g(\pi)$.

The Parseval equality has many very important consequences, some of which we shall describe.

Theorem 5.11 *Two continuous functions on $[-\pi, \pi]$ that have the same Fourier series must coincide.*

PROOF If the functions $f(x)$ and $g(x)$ have the same Fourier series, then all the Fourier coefficients of the function $f(x) - g(x)$ are zero; Parseval's equality then implies that

$$\int_{-\pi}^{\pi} (f(x) - g(x))^2 \, dx = 0,$$

which in turn implies that $f(x) = g(x)$.

A very beautiful consequence of Parseval's equality is the isoperimetric inequality. In order to derive this inequality, we require three lemmas.

Lemma 5.6 *Suppose $f(x)$ is C^1 on $[-\pi, \pi]$, with $f(-\pi) = f(\pi)$, and that $a_0/2 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$ is the Fourier series for $f(x)$. Then the Fourier series for $f'(x)$ is $\sum_{n=1}^{\infty} (nb_n \cos nx - na_n \sin nx)$.*

PROOF This follows immediately from integration by parts, if we compute the Fourier coefficients of $f'(x)$.

Lemma 5.7 (Wirtinger's Inequality) *Suppose $f(x)$ is C^1 on $[-\pi, \pi]$, with $f(-\pi) = f(\pi)$ and $\int_{-\pi}^{\pi} f(x) \, dx = 0$. Then*

$$\int_{-\pi}^{\pi} (f'(x))^2 \, dx \geq \int_{-\pi}^{\pi} (f(x))^2 \, dx.$$

Moreover,

$$\int_{-\pi}^{\pi} (f'(x))^2 \, dx = \int_{-\pi}^{\pi} (f(x))^2 \, dx$$

if and only if $f(x)$ is of the form $a \cos x + b \sin x$, for some constants a and b.

PROOF Suppose $\sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$ is the Fourier series for $g(x)$. (Note that the a_0 term is absent, since $\int_{-\pi}^{\pi} f(x) \, dx = 0$.) By Lemma 5.6, the Fourier series for $f'(x)$ is $\sum_{n=1}^{\infty} (nb_n \cos nx - na_n \sin nx)$. By Parseval's equality, $(1/\pi) \int_{-\pi}^{\pi} (f(x))^2 \, dx = \sum_{n=1}^{\infty} a_n^2 + b_n^2$ and

$$\frac{1}{\pi} \int_{-\pi}^{\pi} (f'(x))^2 \, dx = \sum_{n=1}^{\infty} n^2(a_n^2 + b_n^2).$$

That is,

$$\int_{-\pi}^{\pi} (f'(x))^2 \, dx - \int_{-\pi}^{\pi} (f(x))^2 \, dx = \pi \sum_{n=1}^{\infty} (n^2 - 1)(a_n^2 + b_n^2),$$

and the result is clear from this.

Lemma 5.8 *Suppose a and b are two numbers such that $a^2 + b^2 = 1$. Define functions $x(t)$ and $y(t)$ on $[-\pi, \pi]$, by setting*

$$x(t) = -b \cos t + a \sin t \qquad \text{and} \qquad y(t) = a \cos t + b \sin t.$$

Then as t ranges from $-\pi$ to π, the point $(x(t), y(t))$ traces out, in the clockwise sense, the circle of radius 1 about the origin in the x,y plane.

PROOF There exists a number $t_0 \in [-\pi, \pi]$, such that $b = \cos t_0$ and $a = \sin t_0$. Then

$$\begin{aligned} x(t) &= \sin t \sin t_0 - \cos t \cos t_0 \\ &= -\cos (t + t_0), \end{aligned}$$

and

$$\begin{aligned} y(t) &= \sin t \cos t_0 + \cos t \sin t_0 \\ &= \sin (t + t_0). \end{aligned}$$

The result follows immediately from this.

We now return to the isoperimetric inequality. It states, roughly speaking, that among all closed curves of length 2π, the one that encloses the greatest area is the circle of radius 1, or, what is the same thing, for any closed curve of length 2π that encloses a region of area A, we must have the inequality $2A \leq 2\pi$, with equality if and only if the curve is the unit circle. In order to make this statement more precise, it is necessary to first lay down rules about which curves of length 2π will be admitted as candidates in the competition for greatest area. In what follows, we shall restrict our attention to closed, simple (that is, without self-intersection) curves of class C^1. More specifically, we suppose that our curve is traced out by the point $(x(t), y(t))$ as t ranges from $-\pi$ to π, and that the functions $x(t)$ and $y(t)$ are periodic of period 2π, and of class C^1 on $[-\pi, \pi]$. Moreover, we suppose that the curve is traced out in the clockwise sense, and that it is parametrized by arc length, that is, that $(x'(t))^2 + (y'(t))^2 = 1$, for $t \in [-\pi, \pi]$. Finally, we require that

$$\int_{-\pi}^{\pi} x(t)\, dt = \int_{-\pi}^{\pi} y(t)\, dt = 0.$$

If this is not initially the case, it can always be brought about by a translation of the curve, which has no effect on arc length or area.

Now the arc length of the curve, namely 2π, is given by the integral $\int_{-\pi}^{\pi} [(x'(t))^2 + (y'(t))^2]^{1/2}\, dt$, and since

$$[(x'(t))^2 + (y'(t))^2]^{1/2} = (x'(t))^2 + (y'(t))^2 = 1,$$

we can write the last integral as $\int_{-\pi}^{\pi} [(x'(t))^2 + (y'(t))^2]\, dt$.

On the other hand, the area enclosed by the curve is given by the integral $\int_{-\pi}^{\pi} y(t)x'(t)\,dt$ (*cf.* [1], p. 365). Thus, we must show that

$$\int_{-\pi}^{\pi} [(x'(t))^2 + (y'(t))^2]\,dt - 2\int_{-\pi}^{\pi} y(t)x'(t)\,dt \geq 0,$$

with equality if and only if the curve is a circle of radius 1.

Now the quantity on the left of the last inequality can be written

$$\int_{-\pi}^{\pi} (y(t) - x'(t))^2\,dt + \int_{-\pi}^{\pi} (y'(t))^2\,dt - \int_{-\pi}^{\pi} (y(t))^2\,dt.$$

The first of the last three integrals is clearly ≥ 0, since the integrand is ≥ 0, and by Wirtinger's inequality,

$$\int_{-\pi}^{\pi} (y'(t))^2\,dt \geq \int_{-\pi}^{\pi} (y(t))^2\,dt,$$

We have thus shown that $2A \leq 2\pi$, if we denote by A the area of the region enclosed by our curve. It remains to show that if $2A = 2\pi$, then the curve must be a circle of radius 1. Now the statement that $2A = 2\pi$ is equivalent to the equality

$$\int_{-\pi}^{\pi} (y(t) - x'(t))^2\,dt + \int_{-\pi}^{\pi} (y'(t))^2\,dt - \int_{-\pi}^{\pi} (y(t))^2\,dt = 0.$$

By Wirtinger's inequality

$$\int_{-\pi}^{\pi} (y'(t))^2\,dt - \int_{-\pi}^{\pi} (y(t))^2\,dt \geq 0,$$

so we must have

$$\int_{-\pi}^{\pi} (y(t) - x'(t))^2\,dt = 0,$$

and since the integrand is continuous, this implies that $y(t) = x'(t)$, for all $t \in [-\pi, \pi]$,

On the other hand, it must also be the case that

$$\int_{-\pi}^{\pi} (y'(t))^2\,dt = \int_{-\pi}^{\pi} (y(t))^2\,dt,$$

and this implies, again by Wirtinger's inequality, that $y(t)$ must be of the form $a \cos t + b \sin t$. Thus, since $x'(t) = y(t)$, and $\int_{-\pi}^{\pi} x(t)\,dt = 0$, we conclude that $x(t) = -b \cos t + a \sin t$.

Finally, since

$$(x'(t))^2 + (y'(t))^2 = (a \cos t + b \sin t)^2 + (-b \cos t + a \sin t)^2 = a^2 + b^2 = 1,$$

we conclude, by Lemma 5.8, that the curve in question must be the unit circle about the origin, and this completes the proof of the isoperimetric inequality.

5 Equidistribution

A very beautiful topic, upon which we shall touch briefly here, is provided by the theory of equidistributed sequences.

Suppose $a_1, a_2, \ldots$ is a sequence of numbers tending to ∞. For each n, denote by $\bar{a}_n$ the fractional part of a_n, that is, the difference between a_n and the largest integer less than or equal to a_n. Then $\bar{a}_n \in [0, 1]$ for each n, and it is natural to expect that, in general, the numbers $\bar{a}_n$ are more or less evenly distributed throughout the interval $[0, 1]$. What exactly should we mean by "evenly distributed," or, as it is more commonly called, "equidistributed"?

Definition 5.1 The sequence $a_1, a_2, \ldots$, with $a_n \to \infty$ as $n \to \infty$, is said to be *equidistributed* mod 1, if, for any subinterval (a, b) of $[0, 1]$, the proportion of the numbers $\bar{a}_1, \ldots, \bar{a}_n$ that fall in (a, b) tends to $b - a$ as $n \to \infty$. In other words, if $g(n)$ denotes the number of the $\bar{a}_j$ $(j = 1, \ldots, n)$ that lie in (a, b), then $g(n)/n \to b - a$ as $n \to \infty$.

Lemma 5.9 *Suppose $a_1, a_2, \ldots$ is a sequence such that $a_n \to \infty$ as $n \to \infty$. Suppose, moreover, that for each continuous function $f(x)$ on $[0, 1]$ which is periodic (i.e., $f(0) = f(1)$), $\lim_{N \to \infty} (1/N) \sum_{n=1}^{N} f(\bar{a}_n)$ exists and equals $\int_0^1 f(x)\, dx$. Then the sequence $a_1, a_2, \ldots$ is equidistributed* mod 1.

PROOF Suppose (a, b) is a subinterval of $[0, 1]$. Let $h(x)$ be the function on $[0, 1]$ that equals 1 for $x \in (a, b)$ and 0 otherwise. Then it is easy to see (Figure 14) that for any $\epsilon > 0$ there exist continuous periodic functions $f_1(x)$ and $f_2(x)$ on $[0, 1]$, such that

1. $f_1(x) \le h(x) \le f_2(x), \qquad x \in [0, 1],$
2. $\left| \int_0^1 f_j(x)\, dx - (b - a) \right| \le \epsilon, \qquad \text{for } j = 1, 2.$

(Figure 14 illustrates the case $a \neq 0$, $b \neq 1$. How could we define $f_2(x)$ if $a = 0$ and/or $b = 1$?)

Now by condition (1),

$$\frac{1}{N} \sum_{n=1}^{N} f_1(\bar{a}_n) \le \frac{1}{N} \sum_{n=1}^{N} h(\bar{a}_n) \le \frac{1}{N} \sum_{n=1}^{N} f_2(\bar{a}_n)$$

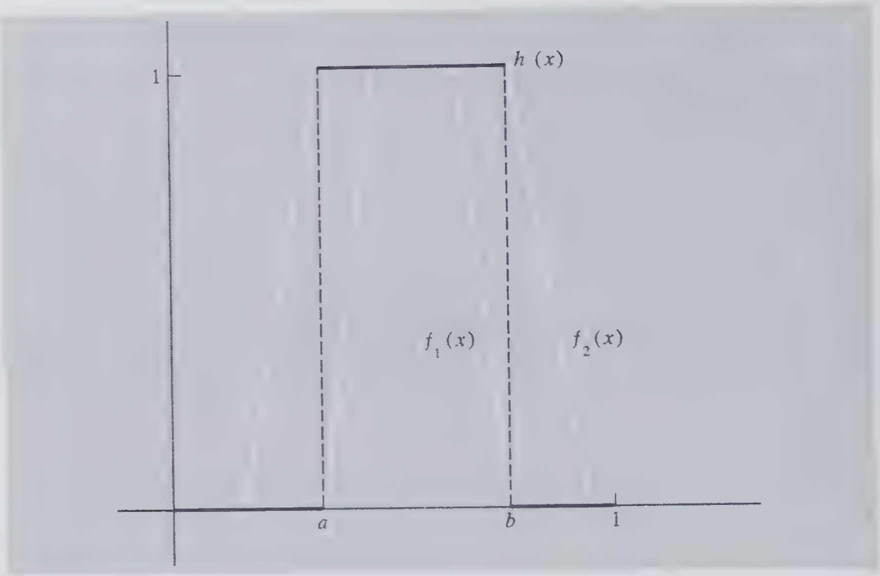

FIGURE 14

for each N. Furthermore, by hypothesis,

$$\frac{1}{N}\sum_{n=1}^{N} f_j(\bar{a}_n) \to \int_0^1 f_j(x)\,dx \qquad \text{as } N\to\infty \quad (j = 1, 2),$$

and since ϵ is arbitrary, condition (2) implies that $(1/N)\sum_{n=1}^{N} h(\bar{a}_n) \to b - a$ as $N \to \infty$, which implies that $a_1, a_2, \ldots$ is equidistributed mod 1.

Lemma 5.10 *Suppose $a_1, a_2, \ldots$ is a sequence such that $a_n \to \infty$, and such that for each positive integer k,*

$$\frac{1}{N}\sum_{n=1}^{N} \cos 2\pi k a_n \to 0 \qquad \textit{and} \qquad \frac{1}{N}\sum_{n=1}^{N} \sin 2\pi k a_n \to 0$$

as $N \to \infty$. Then the sequence $a_1, a_2, \ldots$ is equidistributed mod 1.

PROOF Note, to begin with, that the hypotheses imply that for any trigonometric polynomial $P(x)$ of the form

$$a_0 + \sum_{n=1}^{M} (a_n \cos 2\pi n x + b_n \sin 2\pi n x),$$

$$\frac{1}{N}\sum_{n=1}^{N} P(\bar{a}_n) \to \int_0^1 P(x)\,dx = a_0.$$

By the last lemma, it suffices to prove that for any continuous periodic function $f(x)$ on $[0, 1]$,

$$\frac{1}{N}\sum_{n=1}^{N} f(\bar{a}_n) \to \int_0^1 f(x)\,dx \qquad \text{as } N \to \infty.$$

Now it follows from problem 2 of Chapter 4, that any continuous periodic function on $[0, 1]$ can be uniformly approximated by trigonometric polynomials of the form $a_0 + \sum_{n=1}^{M} (a_n \cos 2\pi nx + b_n \sin 2\pi nx)$. In particular, therefore, by approximating $f(x) - \epsilon$ and $f(x) + \epsilon$ for small $\epsilon > 0$, we can find trigonometric polynomials $P_1(x)$ and $P_2(x)$ of this type, such that

1. $P_1(x) \le f(x) \le P_2(x)$, $x \in [0, 1]$.
2. $\left|\int_0^1 P_j(x)\,dx - \int_0^1 f(x)\,dx\right|$ is as small as we please, for $j = 1, 2$.

Thus, since $(1/N)\sum_{n=1}^{N} P_1(\bar{a}_n) \le (1/N)\sum_{n=1}^{N} f(\bar{a}_n) \le (1/N)\sum_{n=1}^{N} P_2(\bar{a}_n)$ for each N, and since $(1/N)\sum_{n=1}^{N} P_j(x) \to \int_0^1 P_j(x)\,dx$, for $j = 1, 2$, we conclude that $(1/N)\sum_{n=1}^{N} f(\bar{a}_n) \to \int_0^1 f(x)\,dx$, which proves the lemma.

The criterion for equidistribution furnished by the last lemma may appear at first sight to be unwieldy and difficult to apply. In many interesting cases, this is not at all true. We illustrate this with a particularly simple, but nonetheless important, case.

Theorem 5.12 *Suppose α is an irrational number. Then the sequence $\alpha, 2\alpha, 3\alpha, \ldots$ is equidistributed* mod 1.

PROOF By the last lemma, it suffices to show that $(1/N)\sum_{n=1}^{N} \cos 2\pi kn\alpha \to 0$ and $(1/N)\sum_{n=1}^{N} \sin 2\pi kn\alpha \to 0$ as $N \to \infty$, for each positive integer k. Now it is a very simple matter to show, by methods similar to those used in the proof of Lemma 5.5, that

$$\frac{1}{N}\sum_{n=1}^{N} \cos 2\pi kn\alpha = \frac{\sin N\pi k\alpha \cos (N+1)\pi k\alpha}{N \sin \pi k\alpha},$$

and

$$\frac{1}{N}\sum_{n=1}^{N} \sin 2\pi kn\alpha = \frac{\sin N\pi k\alpha \sin (N+1)\pi k\alpha}{N \sin \pi k\alpha}.$$

Now if α is irrational, $\sin \pi k\alpha \neq 0$, so as $N \to \infty$, it is clear that both of the above sums tend to zero, which proves the theorem.

We mention without proof the following result, which is considerably deeper than Theorem 5.12, and contains the latter as a particular case.

Theorem 5.13 **(Weyl)** *Suppose* $P(x) = c_0 + c_1x + \cdots + c_kx^k$ *is a polynomial, for which at least one of the coefficients other than* c_0 *is irrational. Then the sequence* $P(1), P(2), P(3), \ldots$ *is equidistributed* mod 1.

We also mention without proof the following very deep result of Vinogradoff.

Theorem 5.14 *Suppose* α *is an irrational number. Then the sequence* $2\alpha, 3\alpha, 5\alpha, 7\alpha, 11\alpha, 13\alpha, \ldots$ *of prime multiples of* α *is equidistributed* mod 1.

Exercises

1 Show that if $f(x)$ is an even function on $[-\pi, \pi]$, then the b_n's in its Fourier series are all zero, while if $f(x)$ is odd, the a_n's are all zero.

2 Is the system $\sin x, \sin 2x, \sin 3x, \ldots$ complete over $[-\pi, \pi]$? Over $[0, \pi]$?

3 Find the Fourier series for the function $\pi - |x|$ on $[-\pi, \pi]$. Use the result to evaluate the sum $\sum_{n \text{ odd}} n^{-2}$.

4 Show that $\sum_{n=1}^{\infty} n^{-2} = (\sum_{n=0}^{\infty} 4^{-n})(\sum_{n \text{ odd}} n^{-2})$, and use this fact, together with the result of problem 3, to evaluate $\sum_{n=1}^{\infty} n^{-2}$.

5 Suppose θ is a real number. Find the Fourier series for $\cos \theta x$. Use the result to find an expression for $\cot \pi\theta$.

6 Find the third-degree polynomial $P(x)$, for which $\int_{-1}^{1} (x^4 - P(x))^2\, dx$ is a minimum.

7 Suppose $f(x)$ is Riemann integrable over $[-\pi, \pi]$, with $f(-\pi) = f(\pi)$, and suppose that $f(x)$ is C^2 on $[-\pi, 0)$ and $(0, \pi]$. Suppose, moreover, that $f(x)$ can be "symmetrized" smoothly, in the sense that $\frac{1}{2}[f(x) + f(-x)]$ is C^2 on $[-\pi, \pi]$. Show, by considering the Fourier series for $\frac{1}{2}[f(x) + f(-x)]$, that the Fourier series for $f(x)$ converges to $f(x)$ at the point $x = 0$. ($f(x)$ can certainly be smoothly "symmetrized" if $f(x)$ and its first two derivatives approach definite limits as x approaches the origin from the left and right, respectively, and if $f(0)$ is the average of the two limiting values of $f(x)$.)

CHAPTER SIX

The Lebesgue Integral

1 The Lebesgue integral

The problem of extending the integral to a larger class than the Riemann integrable functions is of great intrinsic interest, and progress made in this direction has considerably deepened several branches of mathematics.

In this concluding chapter, we present a very brief outline of the theory of the Lebesgue integral. Since this topic could easily fill a book twice the size of this one, our account is of necessity quite condensed. We shall, nevertheless, cover many of the major points of the theory. There are many approaches to the subject; ours is based, for the most part, on the treatment of integration in Loomis [4].

To begin with, we make some comments on notation. Throughout this chapter, we shall be dealing with functions defined on a fixed closed interval $[a, b]$. If we wish to state that an increasing sequence of functions $f_1(x) \leq f_2(x) \leq f_3(x) \leq \cdots$ converges to a function $f(x)$ on $[a, b]$, we say that $f_n(x) \uparrow f(x)$ or $f(x) = \sup_n f_n(x)$. Similarly, if a decreasing sequence $f_1(x) \geq f_2(x) \geq f_3(x) \geq \cdots$ converges to $f(x)$, we indicate this by saying that $f_n(x) \downarrow f(x)$ or $f(x) = \inf_n f_n(x)$. If $f_1(x), \ldots, f_n(x)$ is a finite collection of functions on $[a, b]$, we define the functions $\max (f_1(x), \ldots, f_n(x))$ and $\min (f_1(x), \ldots, f_n(x))$ to be, respectively, the functions whose values at the point x_0 are

$$\max (f_1(x_0), \ldots, f_n(x_0)) \quad \text{and} \quad \min (f_1(x_0), \ldots, f_n(x_0)).$$

Lemma 6.1 (Dini) *Suppose $f_1(x), f_2(x), f_3(x), \ldots$ is a sequence of continuous functions on $[a, b]$, such that $f_n(x) \downarrow 0$. Then $f_n(x) \downarrow 0$ uniformly on $[a, b]$.*

PROOF If the lemma were false, there would exist a number $\epsilon > 0$ and a sequence $x_1, x_2, x_3, \ldots$ of points in $[a, b]$, such that $f_n(x_n) > \epsilon$. By Theorem 2.3, the sequence $x_1, x_2, x_3, \ldots$ has a subsequence $x'_1, x'_2, x'_3, \ldots$ that converges to a point x_0 in $[a, b]$. Now $f_n(x_0) \downarrow 0$, by hypothesis, so there exists some integer n_0 such that $f_{n_0}(x_0) < \epsilon$. By the continuity of $f_{n_0}(x)$, there must exist an integer n_1, such that for $n \geq n_1$, $f_{n_0}(x'_n) < \epsilon$. Since the sequence of functions $f_1(x), f_2(x), f_3(x), \ldots$ is decreasing, this implies that for

$$n \geq \max(n_0, n_1),$$

$f_n(x'_n) < \epsilon$, which is a contradiction. Hence, we conclude that $f_1(x), f_2(x), \ldots$ must tend to zero uniformly on $[a, b]$.

We now perform the initial extension of the integral, and for this purpose, it is convenient to enlarge our definition of a function on $[a, b]$.

Definition 6.1 A function on $[a, b]$, with values in the extended real numbers, is a rule by which there corresponds to each point in $[a, b]$, either a real number or one of the two symbols ∞ or $-\infty$.

With respect to the two symbols ∞ and $-\infty$, we adopt the following conventions:

1. For any real number c,
$$-\infty \leq \infty, \quad c \leq \infty, \quad \text{and} \quad -\infty \leq c.$$
2. For any real number c,
$$\infty + c = c + \infty = \infty + \infty = \infty,$$
$$-\infty + c = c + (-\infty) = -\infty + (-\infty) = -\infty.$$
3. If c is a positive real number,
$$c \times \infty = \infty \times c = \infty \quad \text{and} \quad c \times (-\infty) = -\infty \times c = -\infty.$$
If c is a negative real number, then
$$c \times \infty = \infty \times c = -\infty \quad \text{and} \quad c \times (-\infty) = -\infty \times c = \infty.$$
$$0 \times \infty = \infty \times 0 = 0 \quad \text{and} \quad 0 \times (-\infty) = -\infty \times 0 = 0.$$
4. If $\{s_n\}$ is a sequence in which each term is either an ordinary real number, or one of the symbols ∞ or $-\infty$, we say that $\sup s_n = \infty$, if there is no ordinary real number that is $\geq$ to all the s_n's. If there is no ordinary real number which is $\leq$ to all the s_n's, we say that
$$\inf_n s_n = -\infty.$$

In the following, we use the word "function" to mean "function on $[a, b]$ with values in the extended real numbers." This, of course, includes as a special case our previous definition of a function on $[a, b]$. If a function does not take on either of the values ∞ or $-\infty$, we shall say that it is *finite valued*, and in general, the phrase "finite number" means an ordinary real number.

We now define a class of functions that is important for the development of the theory of the integral.

Definition 6.2 Define A to be the *class of functions of the form* $\sup_n f_n(x)$, where $f_1(x) \le f_2(x) \le f_3(x) \le \cdots$ is an increasing sequence of continuous functions.

Lemma 6.2 *Suppose $f(x)$ and $g(x)$ are in A and $f(x) \ge g(x)$. Suppose, moreover, that $f_1(x) \le f_2(x) \le f_3(x) \le \cdots$ and $g_1(x) \le g_2(x) \le g_3(x) \le \cdots$ are two increasing sequences of continuous functions, such that $f_n(x) \uparrow f(x)$ and $g_n(x) \uparrow g(x)$. Then $\sup_n \int_a^b f_n(x)\,dx \ge \sup_n \int_a^b g_n(x)\,dx$.*

PROOF Suppose k is a positive integer. For each positive integer n, define a continuous function $h_n(x)$ by setting $h_n(x) = \min(f_n(x), g_k(x))$. Then clearly $f_n(x) \ge h_n(x)$, for each n, and since $f_n(x) \uparrow f(x)$ and $f(x) \ge g_k(x)$, it is clear that $h_n(x) \uparrow g_k(x)$. This implies that as $n \to \infty$, $g_k(x) - h_n(x) \downarrow 0$, so by Lemma 6.1, $h_n(x) \uparrow g_k(x)$ uniformly. By Theorem 2.9, this implies that $\int_a^b h_n(x)\,dx \to \int_a^b g_k(x)\,dx$. Since $f_n(x) \ge h_n(x)$, this in turn implies that $\sup_n \int_a^b f_n(x)\,dx \ge \int_a^b g_k(x)\,dx$, and this implies that

$$\sup_n \int_a^b f_n(x)\,dx \ge \sup_n \int_a^b g_n(x)\,dx,$$

which proves the lemma.

Corollary If $f(x)$ is in A, and if $f_1(x) \le f_2(x) \le f_3(x) \le \cdots$ and $f_1^*(x) \le f_2^*(x) \le f_3^*(x) \le \cdots$ are two increasing sequences of continuous functions, such that $\sup_n f_n(x) = \sup f_n^*(x) = f(x)$, then

$$\sup_n \int_a^b f_n(x)\,dx = \sup_n \int_a^b f_n^*(x)\,dx.$$

Definition 6.3 Suppose $f(x)$ is in A, and suppose $f_1(x) \le f_2(x) \le \cdots$ is an increasing sequence of continuous functions, such that $f_n(x) \uparrow f(x)$. [Such a sequence exists, by the definition of A.] We define the *integral* of $f(x)$,

written $\int_a^b f(x)\,dx$, to be $\sup_n \int_a^b f_n(x)\,dx$. By the Corollary to Lemma 6.2, this definition is unambiguous, since the value of the integral does not depend on the particular sequence $\{f_n(x)\}$. Note that the integral of a function in A could very well have the value ∞.

We now list some important properties of the class A.

1. If $f(x)$ is in A, then so is $cf(x)$, for any finite constant $c \geq 0$. Moreover,

$$\int_a^b (cf(x))\,dx = c\int_a^b f(x)\,dx.$$

Proof: If $f_n(x) \uparrow f(x)$, then $cf_n(x) \uparrow cf(x)$ and

$$\sup_n \int_a^b (cf_n(x))\,dx = c \sup_n \int_a^b f_n(x)\,dx.$$

2. If $f(x)$ and $g(x)$ are in A and if $f(x) \geq g(x)$, then

$$\int_a^b f(x)\,dx \geq \int_a^b g(x)\,dx.$$

Proof: This follows immediately from Lemma 6.2.

3. If $f(x)$ and $g(x)$ are in A, then so are

$$\max(f(x), g(x)) \quad \text{and} \quad \min(f(x), g(x)).$$

Proof: If $f_n(x) \uparrow f(x)$ and $g_n(x) \uparrow g(x)$, then

$$\max(f_n(x), g_n(x)) \uparrow \max(f(x), g(x))$$

and

$$\min(f_n(x), g_n(x)) \uparrow \min(f(x), g(x)).$$

By an obvious induction, it follows that if $f_1(x), \ldots, f_k(x)$ are in A, then so are $\max(f_1(x), \ldots, f_k(x))$ and $\min(f_1(x), \ldots, f_k(x))$.

4. If $f(x)$ and $g(x)$ are in A, then so is $f(x) + g(x)$ and

$$\int_a^b (f(x) + g(x))\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx.$$

Proof: If $f_n(x) \uparrow f(x)$ and $g_n(x) \uparrow g(x)$, then

$$f_n(x) + g_n(x) \uparrow f(x) + g(x).$$

Moreover, for any positive integer n,

$$\int_a^b (f_n(x) + g_n(x))\,dx = \int_a^b f_n(x)\,dx + \int_a^b g_n(x)\,dx,$$

and it follows from this that

$$\int_a^b (f_n(x) + g_n(x))\,dx \le \int_a^b f(x)\,dx + \int_a^b g(x)\,dx,$$

which implies that

$$\int_a^b (f(x) + g(x))\,dx \le \int_a^b f(x)\,dx + \int_a^b g(x)\,dx.$$

On the other hand, it follows from the fact that the sequences $\{f_n(x)\}$ and $\{g_n(x)\}$ are increasing, that for any positive integers n_1 and n_2,

$$\int_a^b (f(x) + g(x))\,dx \ge \int_a^b f_{n_1}(x)\,dx + \int_a^b g_{n_2}(x)\,dx,$$

and this implies that

$$\int_a^b (f(x) + g(x))\,dx \ge \int_a^b f(x)\,dx + \int_a^b g(x)\,dx,$$

that is,

$$\int_a^b (f(x) + g(x))\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx,$$

which proves property (4).

The next theorem shows that the class A cannot be enlarged by taking limits of increasing sequences of functions in A.

Theorem 6.1 *Suppose $f_1(x) \le f_2(x) \le f_3(x) \le \cdots$ is an increasing sequence of functions in A. Define $f(x) = \sup\limits_n f_n(x)$. Then $f(x)$ is in A. Moreover, $\int_a^b f_n(x)\,dx \uparrow \int_a^b f(x)\,dx$.*

PROOF For each $f_n(x)$ there is an increasing sequence $h_{n1}(x) \le h_{n2}(x) \le h_{n3}(x) \le \cdots$ of continuous functions, such that $f_n(x) = \sup\limits_m h_{nm}(x)$.

Now suppose we enumerate, in some way, the functions in the array

$$\begin{array}{cccc}
h_{11}(x) & h_{12}(x) & h_{13}(x) & \cdots \\
h_{21}(x) & h_{22}(x) & h_{23}(x) & \cdots \\
h_{31}(x) & h_{32}(x) & h_{33}(x) & \cdots \\
\vdots & \vdots & \vdots & \cdots
\end{array}$$

As we have seen in Chapter 1, there are many ways of doing this. Let $g_n(x)$ be the nth function in the enumeration that has been selected. Define an increasing sequence of continuous functions $G_1(x) \le G_2(x) \le G_3(x) \le \cdots$ by setting $G_n(x) = \max(g_1(x), \ldots, g_n(x))$. Then it is clear that for all n, $G_n(x) \le f(x)$. Thus, the function $G(x)$ in A, defined by setting

$$G(x) = \sup_n G_n(x),$$

satisfies $G(x) \le f(x)$. On the other hand, it is evident that for any function $h_{ij}(x)$, $h_{ij}(x) \le G(x)$. In particular, it must therefore be the case that for each n, $\sup_m h_{nm}(x) \le G(x)$ or, what is the same thing, $f_n(x) \le G(x)$. Since $f(x) = \sup_n f_n(x)$, this implies that $f(x) \le G(x)$. That is, we have shown that $f(x) \le G(x) \le f(x)$, which implies that $f(x) = G(x)$, and this shows that $f(x)$ is in A. In order to show that $\int_a^b f_n(x)\,dx \uparrow \int_a^b f(x)\,dx$, note, to begin with, that $\sup_n \int_a^b f_n(x)\,dx \le \int_a^b f(x)\,dx$. On the other hand, it is clear that for any n, there exists an m, such that $G_n(x) \le f_m(x)$. This implies that

$$\sup_n \int_a^b G_n(x)\,dx \le \sup_n \int_a^b f_n(x)\,dx,$$

and since $\int_a^b f(x)\,dx = \sup_n \int_a^b G_n(x)\,dx$, this in turn implies that

$$\int_a^b f(x)\,dx \le \sup_n \int_a^b f_n(x)\,dx.$$

That is, $\int_a^b f(x)\,dx \le \sup_n \int_a^b f_n(x)\,dx \le \int_a^b f(x)\,dx$, and this implies that $\sup_n \int_a^b f(x)\,dx = \int_a^b f(x)\,dx$, or what is the same thing, that

$$\int_a^b f_n(x)\,dx \uparrow \int_a^b f(x)\,dx,$$

and this concludes the proof of the theorem.

Definition 6.4 By the class $-A$ we mean the *class of functions of the form* $-f(x)$, where $f(x)$ is in A. Equivalently, $-A$ consists of all functions of the form $\inf_n f_n(x)$, where $f_1(x) \ge f_2(x) \ge f_3(x) \ge \cdots$ is a decreasing sequence of continuous functions. For $f(x)$ in $-A$, we define the integral of $f(x)$ by setting $\int_a^b f(x)\,dx = -\int_a^b (-f(x))\,dx$.

Note that if $f(x)$ is in A and $g(x)$ is in $-A$, and if, moreover, $f(x) \ge g(x)$, then $\int_a^b f(x)\,dx \ge \int_a^b g(x)\,dx$. In order to see this, note that by definition the

function $-g(x)$ is in A, and the statement that $f(x) \geq g(x)$ is equivalent to the statement that $f(x) + (-g(x)) \geq 0$. From this it follows that

$$\int_a^b f(x)\,dx + \int_a^b (-g(x))\,dx \geq 0$$

or, what is the same thing,

$$\int_a^b f(x)\,dx \geq -\int_a^b (-g(x))\,dx,$$

which is equivalent to the assertion.

We now extend the integral to a still larger class of functions. The definition of this new class of functions is similar to the definition of the class of Riemann integrable functions, except that here the role of the step functions is played by appropriate functions in A and $-A$.

Definition 6.5 We say that a function $f(x)$ is in the class B if, for any $\epsilon > 0$, there exist functions $h(x)$ and $g(x)$ in $-A$ and A, respectively, both having finite integrals, and such that

1. $h(x) \leq f(x) \leq g(x)$
2. $\int_a^b g(x)\,dx - \int_a^b h(x)\,dx \leq \epsilon.$

If $f(x)$ is in B, we define the integral of $f(x)$ by setting

$$\int_a^b f(x)\,dx = \sup_{\substack{h(x)\in -A\\ h(x)\leq f(x)}} \int_a^b h(x)\,dx.$$

Here $h(x)$ ranges, as indicated, over all functions in $-A$ that are $\leq f(x)$. Equivalently, it is clear from Properties (1) and (2) that we could define

$$\int_a^b f(x)\,dx = \inf_{\substack{g(x)\in A\\ f(x)\leq g(x)}} \int_a^b g(x)\,dx.$$

Remark It is evident from the definition of the class B, that any function in A or $-A$ that has a finite integral must be in B, and must have the same integral, regarded as a function in B. Suppose, for example, $f(x)$ is in A and $\epsilon > 0$ is given. Then by the definition of the class A, there must exist a continuous function $h(x)$, such that $h(x) \leq f(x)$ and $\int_a^b f(x)\,dx - \int_a^b h(x)\,dx \leq \epsilon$. Since any continuous function is in $-A$, this shows that $f(x)$ is in B, if we take the function $g(x)$, which occurs in the definition of the class B, to be $f(x)$ itself. It is, moreover, obvious that the value of the integral of $f(x)$,

regarded as a function in B, must coincide with its value when $f(x)$ is regarded as a function in A, since the function $\inf_{\substack{g(x)\in A \\ f(x)\le g(x)}} g(x)$ equals $f(x)$.

The following properties of B follow easily from its definition.

1. If $f(x)$ is in B, then so is $cf(x)$, for any finite constant c. Moreover, $\int_a^b (cf(x))\,dx = c\int_a^b f(x)\,dx$. (Note that there is no restriction on the signature of c.)

   Proof: If $c = 0$, the assertion is obvious. If $c > 0$ and if $\epsilon > 0$ is given, then there exist, by hypothesis, functions $h(x)$ in $-A$ and $g(x)$ in A, such that $h(x) \le f(x) \le g(x)$, and $\int_a^b g(x)\,dx - \int_a^b h(x)\,dx \le \epsilon/c$. It is then obvious that $ch(x)$ is in $-A$, $cg(x)$ is in A, $ch(x) \le cf(x) \le cg(x)$, and $\int_a^b (cg(x))\,dx - \int_a^b (ch(x))\,dx \le \epsilon$. If, finally, $c < 0$, then $cg(x)$ is in $-A$, $ch(x)$ is in A, $cg(x) \le cf(x) \le ch(x)$, and $\int_a^b (ch(x))\,dx - \int_a^b (cg(x))\,dx \le \epsilon$.

2. If $f(x)$ and $h(x)$ are in B, and if $f(x) \ge g(x)$, then

$$\int_a^b f(x)\,dx \ge \int_a^b g(x)\,dx.$$

   Proof: It is obvious that

$$\sup_{\substack{h(x)\in -A \\ h(x)\le f(x)}} \int_a^b h(x)\,dx \ge \sup_{\substack{h(x)\in -A \\ h(x)\le g(x)}} \int_a^b h(x)\,dx.$$

3. If $f_1(x)$ and $f_2(x)$ are in B, then so are $\max(f_1(x), f_2(x))$ and $\min(f_1(x), f_2(x))$.

   Proof: Suppose $\epsilon > 0$ is given. By hypothesis, there exist functions $h_1(x)$, $h_2(x)$ in $-A$ and functions $g_1(x)$, $g_2(x)$ in A, such that $h_1(x) \le f_1(x) \le g_1(x)$, $h_2(x) \le f_2(x) \le g_2(x)$,

$$\int_a^b g_1(x)\,dx - \int_a^b h_1(x)\,dx \le \tfrac{1}{2}\epsilon$$

   and

$$\int_a^b g_2(x)\,dx - \int_a^b h_2(x)\,dx \le \tfrac{1}{2}\epsilon.$$

   Now it is clear that $\max(h_1(x), h_2(x)) \le \max(f_1(x), f_2(x)) \le \max(g_1(x), g_2(x))$. Moreover, the functions $-h_1(x)$, $-h_2(x)$ and

$-\max(h_1(x), h_2(x))$ are all in A, so the functions $\max(g_1(x), g_2(x)) + (-\max(h_1(x), h_2(x))$, $g_1(x) + (-h_1(x))$, and $g_2(x) + (-h_2(x))$ are also in A. It is, moreover, evident that

$$\max(g_1(x), g_2(x)) + (-\max(h_1(x), h_2(x))) \\ \leq (g_1(x) + (-h_1(x))) + (g_2(x) + (-h_2(x))),$$

and this implies that

$$\int_a^b \max(g_1(x), g_2(x))\,dx - \int_a^b \max(h_1(x), h_2(x))\,dx \leq \tfrac{1}{2}\epsilon + \tfrac{1}{2}\epsilon = \epsilon.$$

An analogous argument shows that $\min(f_1(x), f_2(x))$ is in B. By an obvious induction, it follows that if $f_1(x), \ldots, f_n(x)$ are in B, then so are $\max(f_1(x), \ldots, f_n(x))$ and $\min(f_1(x), \ldots, f_n(x))$.

Remark We have not said anything about sums of functions in B. The reason for this is that we wish to avoid the essentially ambiguous symbol $\infty + (-\infty)$. We avoid this difficulty by passing to a slightly restricted version of the class B, and in so doing, we arrive, finally, at our goal, namely the Lebesgue class $\mathrm{L}^1[a, b]$.

Definition 6.6 We say that a function $f(x)$ is Lebesgue integrable, if $f(x)$ is a finite-valued function in the class B. The integral of $f(x)$ is defined to be its integral, regarded as a function in B. The class of Lebesgue integrable functions (over the interval $[a, b]$) is denoted by the symbol $\mathrm{L}^1[a, b]$.

Remark It is evident from the definition, that properties (1)–(3) for functions of B hold word for word in $\mathrm{L}^1[a, b]$. It is, moreover, evident that if $f_1(x)$ and $f_2(x)$ are in $\mathrm{L}^1[a, b]$, then so is $f_1(x) + f_2(x)$, and

$$\int_a^b (f_1(x) + f_2(x))\,dx = \int_a^b f_1(x)\,dx + \int_a^b f_2(x)\,dx.$$

(The same is, of course, true for differences, by property (1) of the class B, taking $c = -1$.) The proof of the additivity of the Lebesgue integral is very similar to the proof of property (3) for class B, and is omitted.

Note that if $f(x)$ is in $\mathrm{L}^1[a, b]$, then so is $|f(x)|$, since

$$|f(x)| = \max(f(x), 0) - \min(f(x), 0).$$

Remark It is easy to see that any function that is Riemann integrable over $[a, b]$ must be in $\mathrm{L}^1[a, b]$ and have the same integral when regarded as a

function in $L^1[a, b]$. In order to show this, it certainly suffices to show that any step-function $f(x)$ is in B, and that its integral, when $f(x)$ is regarded as a function in B, coincides with its Riemann integral. To show this, it in turn suffices to consider the special case in which $f(x) = 1$, for x in some subinterval I of $[a, b]$, and $f(x) = 0$ otherwise, since any step-function is a linear combination of such functions. Figure 15 illustrates how this can be done.

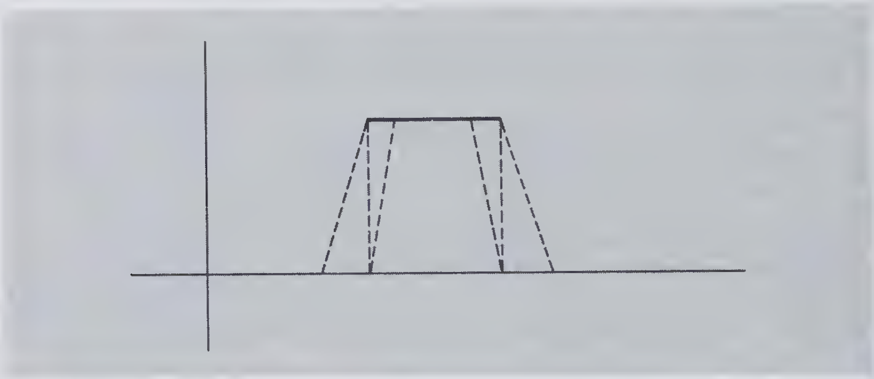

FIGURE 15

The following convergence theorem is very useful.

Theorem 6.2 **(Beppo Levi)** *Suppose* $f_1(x) \leq f_2(x) \leq f_3(x) \leq \cdots$ *is an increasing sequence of functions in* $L^1[a, b]$ *Suppose, moreover, that the associated sequence of integrals* $\int_a^b f_1(x)\,dx, \int_a^b f_2(x)\,dx, \ldots$ *is bounded above. Then the function* $f(x) = \sup_n f_n(x)$ *is in class B. Moreover,*

$$\int_a^b f_n(x)\,dx \uparrow \int_a^b f(x)\,dx.$$

PROOF It is clear that in proving this theorem we may suppose that $f_1(x) \equiv 0$; for if this is not initially the case, it can be brought about by replacing the sequence $\{f_n(x)\}$ by the sequence $\{f_n(x) - f_1(x)\}$, and it is evident that if the theorem is true for the latter sequence, then it must be true for the original sequence. Accordingly, suppose $f_1(x) \equiv 0$. Let $\epsilon > 0$ be given, and for each integer $n \geq 2$, let $g_n(x)$ be a function in A, such that $f_n(x) - f_{n-1}(x) \leq g_n(x)$ and $\int_a^b g_n(x)\,dx - \int_a^b (f_n(x) - f_{n-1}(x))\,dx \leq \epsilon/2^n$, or equivalently, $\int_a^b g_n(x)\,dx \leq \int_a^b (f_n(x) - f_{n-1}(x))\,dx + \epsilon/2^n$. Note that each function $g_n(x)$ must be ≥ 0, since $f_n(x) - f_{n-1}(x) \geq 0$, so the sequence of functions $g_2(x)$, $g_2(x) + g_3(x)$, $g_2(x) + g_3(x) + g_4(x), \ldots$ is increasing, and hence the function $g(x) = \sup_n (\sum_{k=2}^{\infty} g_k(x))$ must be in A. Moreover,

$\int_a^b g(x)\,dx = \sup_n \int_a^b (\sum_{k=2}^{\infty} g_k(x))\,dx$ by Theorem 6.1. Now for each $n \geq 2$, $f_n(x) = (f_2(x) - f_1(x)) + \cdots + (f_n(x) - f_{n-1}(x))$, since everything on the right side cancels out except $f_n(x)$ and $-f_1(x)$, but $-f_1(x) \equiv 0$. From this we conclude that for each $n \geq 2$, $f_n(x) \leq \sum_{k=2}^{n} g_k(x)$, and hence $f(x) \leq g(x)$. On the other hand, for each $n \geq 2$, $\int_a^b f_n(x)\,dx = \int_a^b (f_2(x) - f_1(x))\,dx + \cdots + \int_a^b (f_n(x) - f_{n-1}(x))\,dx$, so $\int_a^b (\sum_{k=2}^{n} g_k(x))\,dx \leq \int_a^b f_n(x)\,dx + \epsilon/2^2 + \cdots + \epsilon/2^n \leq \int_a^b f_n(x)\,dx + \frac{1}{2}\epsilon$.

From this it follows that $\int_a^b g(x)\,dx \leq \sup_n \int_a^b f_n(x)\,dx + \frac{1}{2}\epsilon$, or equivalently, $\int_a^b g(x)\,dx - \sup_n \int_a^b f_n(x)\,dx \leq \frac{1}{2}\epsilon$. Now there clearly exists an integer n_0, such that $\sup_n \int_a^b f_n(x)\,dx - \int_a^b f_{n_0}(x)\,dx \leq \frac{1}{4}\epsilon$, and since $f_{n_0}(x)$ is in $L^1[a, b]$, there exists a function $h(x)$ in $-A$, having a finite integral, such that $h(x) \leq f_{n_0}(x)$ and $\int_a^b f_{n_0}(x)\,dx - \int_a^b h(x)\,dx \leq \frac{1}{4}\epsilon$. From this we conclude that $\sup_n \int_a^b f_n(x)\,dx - \int_a^b h(x)\,dx \leq \frac{1}{2}\epsilon$. Since $\int_a^b g(x)\,dx - \sup_n \int_a^b f_n(x)\,dx \leq \frac{1}{2}\epsilon$, we further conclude that $\int_a^b g(x)\,dx - \int_a^b h(x)\,dx \leq \epsilon$. Thus, since $h(x) \leq f(x) \leq g(x)$, this proves that $f(x)$ is in B, and since ϵ was arbitrary, and $\int_a^b h(x)\,dx \leq \sup_n \int_a^b f_n(x)\,dx \leq \int_a^b g(x)\,dx$, it is clear that $\int_a^b f(x)\,dx = \sup_n \int_a^b f_n(x)\,dx$, which proves Beppo Levi's theorem.

Corollary *Suppose $f_1(x) \geq f_2(x) \geq f_3(x) \geq \cdots$ is a decreasing sequence of function in $L^1[a, b]$. Suppose, moreover, that the associated sequence of integrals $\int_a^b f_1(x)\,dx, \int_a^b f_2(x)\,dx, \int_a^b f_3(x)\,dx, \ldots$ is bounded below. Then the function $\inf_n f_n(x)$ is in the class B. Moreover, $\int_a^b f_n(x)\,dx \downarrow \int_a^b f(x)\,dx$.*

The following important result can be derived very easily from Beppo Levi's theorem.

Theorem 6.3 **(Lebesgue Dominated Convergence Theorem)** *Suppose $f_1(x), f_2(x), f_3(x), \ldots$ is a sequence of functions in $L^1[a, b]$, and suppose there exists a function $F(x)$ in $L^1[a, b]$, such that $|f_n(x)| \leq F(x)$, for all $n \geq 1$. Suppose, finally, that there exists a function $f(x)$, such that $f_n(x) \to f(x)$. Then $f(x)$ must be in $L^1[a, b]$, and $\int_a^b f_n(x)\,dx \to \int_a^b f(x)\,dx$.*

PROOF For each $n \geq 1$, the function $\sup (f_1(x), \ldots, f_n(x))$ is in $L^1[a, b]$ Moreover,

$$f_1(x) \leq \sup (f_1(x), f_2(x)) \leq \sup (f_1(x), f_2(x), f_3(x)) \leq \cdots$$

and

$$\int_a^b (\sup (f_1(x), \ldots, f_n(x)))\, dx \le \int_a^b F(x)\, dx, \qquad \text{for all } n \ge 1.$$

Thus, by Beppo Levi's theorem, the function $g_1(x) = \sup (f_1(x), f_2(x), f_3(x), \ldots)$ is in B, and has a finite integral. Moreover, since it is clear that $|g_1(x)| \le F(x)$, it follows that $g_1(x)$ is finite-valued, and hence in $L^1[a, b]$. By similar reasoning, we see that the functions $g_n(x) = \sup (f_n(x), f_{n+1}(x), f_{n+2}(x), \ldots)$ are all in $L^1[a, b]$, and since $|g_n(x)| \le F(x)$, their integrals are all bounded below. Now $g_1(x) \ge g_2(x) \ge g_3(x) \ge \cdots$ and $g_n(x) \downarrow f(x)$, since $f_n(x) \to f(x)$. By the Corollary to Beppo Levi's theorem, we conclude that $f(x)$ is in B, and $\int_a^b g_n(x)\, dx \downarrow \int_a^b f(x)\, dx$. Since it is clear that $|f(x)| \le F(x)$, it follows that $f(x)$ is in $L^1[a, b]$. Now it is easy to see, by reasoning analogous to the above, that the functions $h_n(x) = \inf (f_n(x), f_{n+1}(x), f_{n+2}(x), \ldots)$ are in $L^1[a, b]$, and $\int_a^b h_n(x)\, dx \uparrow \int_a^b f(x)\, dx$. Since for each $n \ge 1$, $h_n(x) \le f_n(x) \le g_n(x)$, we conclude that $\int_a^b f_n(x)\, dx \to \int_a^b f(x)\, dx$, which proves the theorem.

2 Measurable sets

Some insight into the nature of functions in the class B can be gained if we introduce the concept of a set of measure zero.

Definition 6.7 Suppose S is a subset of $[a, b]$. By the *indicator function* or, as it is often called, the *characteristic function* of the set S, we mean the function $I_S(x)$, whose value is 1 for x in S, and zero elsewhere. We say that the set S is *measurable* if its indicator function $I_S(x)$ is in $L^1[a, b]$. If S is measurable, we define the *measure* of S to be $\int_a^b I_S(x)\, dx$. A subset S of $[a, b]$ is said to be of *measure zero* if it is measurable and has measure zero.

Theorem 6.4 *Suppose $f(x)$ is in A and has a finite integral. Denote by S, the set on which $f(x) = \infty$. Then S is of measure zero.*

PROOF Since $f(x)$ can be expressed as the limit of an increasing sequence of continuous functions, it is clear that $f(x)$ is bounded below, and hence, we may assume, by the addition of a constant if necessary, that $f(x) \ge 0$. If we now denote the indicator function of S by $I_S(x)$, it is then clear that for any integer $n > 0$, $0 \le I_S(x) \le (1/n)f(x)$. Since $\int_a^b (1/n)f(x)\, dx \downarrow 0$ as $n \to \infty$, this immediately implies that $I_S(x)$ is in B, and hence in $L^1[a, b]$, since it is finite-valued, and that $\int_a^b I_S(x)\, dx = 0$, which proves the theorem.

Corollary *Suppose $f(x)$ is in $-A$, and has a finite integral. Denote by S, the set on which $f(x) = -\infty$. Then S is of measure zero.*

Corollary *Suppose $f(x)$ is in B, and denote by S, the set on which $f(x)$ has either the value ∞ or $-\infty$. Then S is of measure zero.*

Corollary *Suppose $f_1(x) \leq f_2(x) \leq f_3(x) \leq \cdots$ is an increasing sequence of functions in $L^1[a, b]$, having collectively bounded integrals. Then $\sup_n f_n(x)$ is finite, with the possible exception of the x's in a set of measure zero.*

PROOF This follows immediately from the second Corollary above, and Beppo Levi's theorem.

Theorem 6.5 *Suppose S is a set of measure zero, and suppose $f(x)$ is a finite-valued function that is zero for x not in S. Then $f(x)$ is in $L^1[a, b]$ and $\int_a^b f(x)\,dx = 0$.*

PROOF We may clearly suppose $f(x) \geq 0$, since if the result is established for such functions, it must be true for the difference of two such functions, which implies the general case. Now by hypothesis, $I_S(x)$, the indicator function of S, is in $L^1[a, b]$ and $\int_a^b I_S(x)\,dx = 0$. From this it follows that for any n, the function $nI_S(x)$ is in $L^1[a, b]$ and has integral zero. Since $I_S(x) \leq 2I_S(x) \leq 3I_S(x) \leq \cdots$, it follows from Beppo Levi's theorem that the function $F(x)$, which has the value ∞ for x in S and zero elsewhere, is in B and has integral zero. Thus, there exists a function $g(x)$ in A, having a finite integral, and such that $0 \leq F(x) \leq g(x)$, and hence $0 \leq f(x) \leq g(x)$. Since the functions $(1/n)g(x)$ are in A, and $\int_a^b (1/n)g(x)\,dx \downarrow 0$ as $n \to \infty$, this implies that $f(x)$ is in $L^1[a, b]$ and has integral zero, which was to be shown.

Remark As the last theorem shows, the behavior of a function in $L^1[a, b]$ on a set of measure zero does not affect its integral. For this reason, it is often convenient to regard two functions in $L^1[a, b]$ as equivalent if they coincide except possibly on a set of measure zero. We may then regard $L^1[a, b]$, not as a collection of functions, but rather as the set of equivalence classes corresponding to the above relationship, and all of the important properties that we have established for $L^1[a, b]$ carry over easily to this more abstract point of view. Indeed, from this point of view, $L^1[a, b]$ could equally well be regarded as consisting of equivalence classes of finite-valued functions that need not even be defined on sets of measure zero, but which, where defined, coincide with functions in our original $L^1[a, b]$.

Exercises

1 Show that a subset of a set of measure zero is a set of measure zero.

2 Show that the union of two sets of measure zero is a set of measure zero.

3 Generalize the result of problem 2, by showing that the union of a sequence $S_1, S_2, \ldots$ of sets of measure zero is a set of measure zero. (HINT: use Beppo Levi's theorem.)

4 Do the irrational numbers in [0, 1] constitute a measurable subset of [0, 1]? If so, what is the measure of this subset?

5 In the evaluation of the integral $\int_0^\infty (\sin x)/x\, dx$ by methods from complex variables, it is, at one stage, necessary to show that $\int_0^\pi e^{-r\sin\theta}\, d\theta \to 0$ as $r \to \infty$. Show that this follows from the Lebesgue dominated convergence theorem (Theorem 6.3).

6 Show that Beppo Levi's theorem (Theorem 6.2) is true for functions in the class B.

7 Suppose $a_1, a_2, \ldots$ is a sequence of non-negative numbers such that $\sum_{n=1}^\infty a_n$ converges. For a real number x, denote by $\|x\|$ the distance of x from the nearest integer. Suppose $0 < \theta < 1$. Show that the series $\sum_{n=1}^\infty a_n\|x\|^{-\theta}$ is, except on a set of measure zero, convergent for $x \in [0, 1]$. (HINT: if we define $\|x\|^{-\theta} = \infty$ for x an integer, it is easy to see that $\|x\|^{-\theta}$ is in the class B on [0, 1]. Show that $\int_0^1 \|nx\|^{-\theta}\, dx$ is independent of n, for $n = 1, 2, \ldots$, and apply the result of problem 6.)

References

[1] Courant, R., and F. John. *Introduction to Calculus and Analysis*, Vol. I. (New York City: Interscience Publishers, Inc., 1965).

[2] Knopp, Konrad. *Theory and Application of Infinite Series* (London: Blackie and Son Limited, 1951).

[3] ———. *Theory of Functions, Parts I and II* (New York City: Dover Publications, Inc., 1945).

[4] Loomis, L. H. *An Introduction to Abstract Harmonic Analysis* (Princeton, N.J.: Van Nostrand Co., Inc., 1953).

[5] Stone, M. H. "The Generalized Weierstrass Approximation Theorem," *Mathematics Magazine* **21** (1948), 167–184, 237–254.

Index

Absolute value, 5
Approximation in the mean, theorem of the best, 74

Beppo Levi theorem, 104–105
Bessel's inequality, 73–74

Cauchy product, 20, 21, 25, 45
Cauchy-Hadamard theorem, 42
Cauchy's double series theorem, 21
Cauchy-Schwarz theorem, 74–75
Continuity
  at a point, 26–27
  C^n functions, 35
  over an interval, 27
  uniform, 30
Convergence
  absolute, 16
  alternating series test, 16
  comparison test, 12–13
  conditional, 16
  integral test, 14–15
  majorization test, 12–13
  rate of, 32
  ratio test, 13
  root test, 13–14
  of sequences, 9
  of sequences of functions, 32–34
  of series, 12
  of a series of functions, 32
  uniform, 32, 33

Dedekind cut, 4
Derivative, 35
Divergence, of series, 12
Dini's theorem, 95–96
Dirichlet kernel, 79, 81

Equidistribution, 91–94

Fourier coefficients, 84, 88
Fubini's theorem, 21–23
Function, 26
  analytic, 42
  characteristic, 106
  of class C^n, 35
  of a complex variable, 56
  continuous, 26–27
  differentiable, 34–35
  discontinuous, 27
  even, 38
  finite valued, 97
  indicator, 106
  inverse, 62
  Lebesgue integrable, 103
  maximum of, 29
  minimum of, 29
  odd, 38
  orthonormal, 72, 73
  oscillation, 30
  periodic, 70
  piecewise C^1, 79
  piecewise C^2, 83
  piecewise linear continuous, 59
  polygonal, 59
  Rademacher, 70
  Riemann integrable, 35

Function *(continued)*
step, 35
strictly decreasing, 62
strictly increasing, 62
symmetrized, 94
Fundamental theorem of calculus, 36–37

Greatest lower bound, 7

Infimum, 8
Integral, 35–36
Lebesgue, 95–107
Intermediate value theorem, 29–30
Interval, 26
of convergence of power series, 42
Irrational number, 4
Isoperimetric inequality, 88, 89–90

Karamata Tauberian theorem, 64–69

Least upper bound, 8
Lebesgue dominated convergence theorem, 105–106
Limit
of a sequence, 9
of a series, 12
Linear combination of functions, 71
Linearly independent functions, 71

Measure, 106
set of measure zero, 106
Minkowski's inequality, 75–76

Order, 2, 5
Ordered sum, 56
Orthonormal systems, 70
complete, 76
Oscillation of a function, 30

Parseval's equality, 77
Partial sums, 7
Polynomial
orthonormalized Legendre, 73
trigonometric, 86
Power series, 32

Radius of convergence, 39
ratio method for determining, 40
Rational numbers, 1–7
Real numbers, 3–8
extended real number system, 96–97
Riemann-Lebesgue lemma, 79

Sequence, 8
convergence, 9
decreasing sequence of functions, 95
equidistributed, 91
increasing sequence of functions, 95
orthonormal, 70
orthonormalizing, 73
Series, 11
binomial, 54
Cauchy product of two, 21
convergence of, 12
convergence tests, 12–16
divergence of, 12
Fourier, 80
of functions, 32
geometric, 8
limit, 12
sum, 12
Stone-Weierstrass theorem, 63
Subsequence, 10
Supremum, 7

Taylor expansion, 54
Taylor series, 52
Taylor's theorem, 51
Triangle inequality, 7

Uniform approximation, 58, 59, 62, 69

Vinogradoff's theorem, 94

Weierstrass approximation theorem, 58–61
Weyl's theorem, 94
Wirtinger's inequality, 88

QA
331.5 Randol
.R33 An introduction to
real analysis

HARCOURT, BRACE & WORLD